

RAPPORT VILLANI

DONNER UN SENS À L'INTELLIGENCE ARTIFICIELLE POUR UNE STRATÉGIE NATIONALE ET EUROPÉENNE



ENA, 22 mai 2018
@yann_bonnet
Membre de la mission Villani



En septembre 2017,
le Premier Ministre a chargé le
député et mathématicien
Cédric Villani
d'une mission parlementaire
sur l'**intelligence artificielle (IA)**

Le Premier ministre a chargé Cédric Villani, mathématicien et député, d'une mission sur l'intelligence artificielle (IA) en septembre 2017.

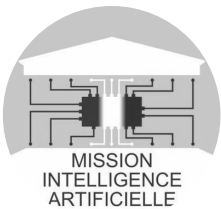


- L'objectif assigné à la mission consiste :
 - à étudier les actions nécessaires pour permettre à la France et à l'Europe d'être à la pointe de l'économie de l'IA,
 - décrire les meilleures pratiques internationales d'application de ces technologies au service de la transformation et de l'amélioration des politiques publiques,
 - identifier les applications prioritaires à déployer à l'intérieur de la sphère publique et
 - ouvrir le champ à une réflexion nationale sur les impacts de l'IA, en considérant ses effets sur le travail et les questions éthiques qu'elle soulève.
- Cette mission intervient après la publication de plusieurs travaux d'importance.
 - Le rapport France IA de mars 2017, comme celui de l'Office parlementaire d'évaluation des choix scientifiques et techniques (OPECST) avant lui, ont permis de lancer le débat en mobilisant la communauté intéressée (acteurs institutionnels, académiques, industriels).
 - La mission devait affiner, approfondir et prioriser ce travail pour le décliner en feuille de route concrète et poser les bases d'une action française et européenne dans la durée.

- Pour mener à bien cette mission, Cédric Villani dispose de l'appui de Marc Schoenauer, directeur de recherche et expert de l'IA à l'INRIA et du secrétariat général du Conseil national du numérique — Yann Bonnet, secrétaire général, Charly Berthet, Anne-Charlotte Cornut et François Levin (rapporteurs au CNNum) et Stella Biabiany assistante de direction — renforcé de Bertrand Rondepierre, ingénieur de l'armement à la DGA.

Méthodologie

une **consultation** en ligne
+ **420** auditions d'experts
un **benchmark** dans 15 pays



Les travaux de la mission se sont déroulés du 8 septembre 2017 au 8 mars 2018.

Ils ont consisté en :

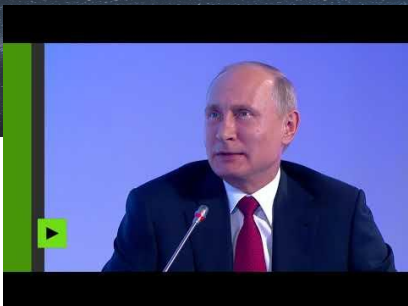
- plus de 400 auditions d'experts issus de domaines variés et concernés par l'intelligence artificielle ;
- la mise en place d'une consultation publique en partenariat avec [Parlement & Citoyens](#), qui a recueilli plus de 2 000 contributions reçues ;
- une étude comparative (benchmark) des politiques mises en place dans 15 pays.



Qu'est ce que l'IA ?

Définir l'intelligence artificielle n'est pas chose facile. Il convient néanmoins de la démystifier.

- Le champ est si vaste qu'il est impossible de la restreindre à un domaine de recherche spécifique.
- C'est plutôt un programme multidisciplinaire.



Pour voir les vidéos, cliquer sur les images

Définir l'IA n'est pas chose facile.

“

Comprendre comment fonctionne la cognition humaine et la reproduire en créant des processus cognitifs comparables à ceux de l'être humain.

”

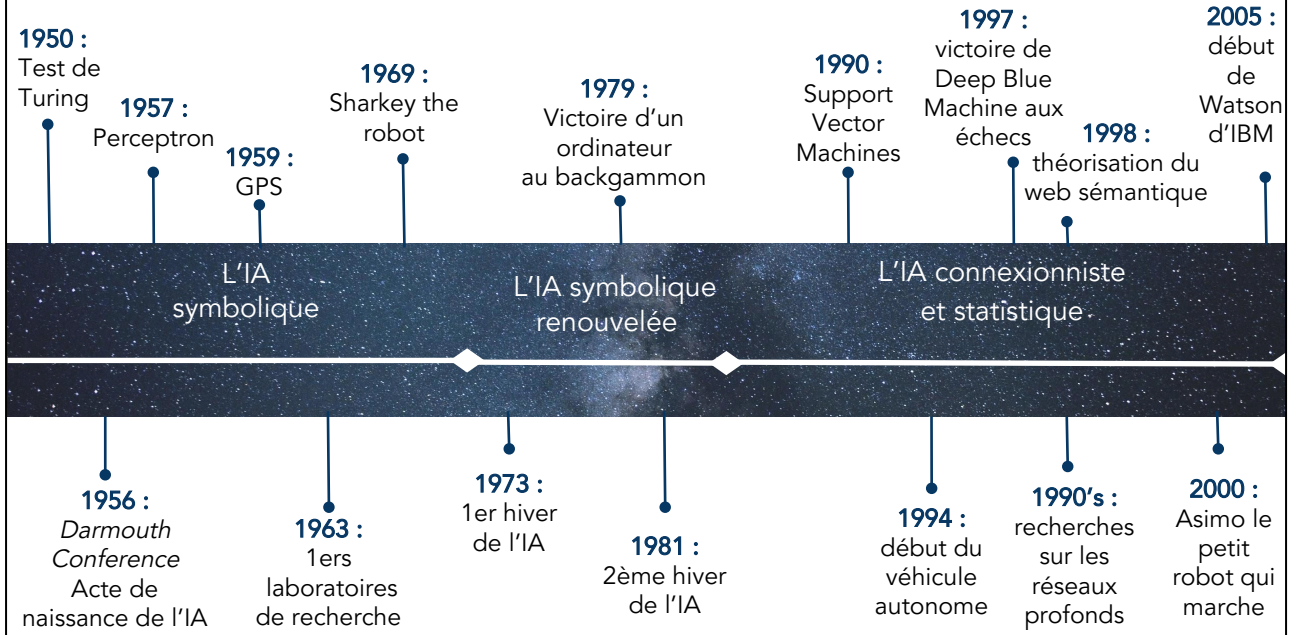
L'ambition initiale de l'IA était d'imiter les processus cognitifs de l'être humain.

Ses objectifs actuels visent plutôt à mettre au point des automates qui résolvent certains problèmes bien mieux que les humains, par tous les moyens disponibles.

- Ainsi l'IA vient au carrefour de plusieurs disciplines : mathématique (logique, optimisation, analyse, probabilités, algèbre linéaire), informatique, sciences cognitives... sans oublier les connaissances spécialisées des domaines auxquelles on souhaite l'appliquer. Et algorithmes qui la sous-tendent reposent sur des approches tout aussi variées : analyse sémantique, représentation symbolique, apprentissage statistique ou exploratoire, réseaux de neurones, etc.
- L'essor récent et foudroyant de l'intelligence artificielle s'explique par les avancées importantes en matière d'apprentissage automatique (machine learning). Les techniques d'apprentissage constituent une révolution par rapport aux approches historiques d'IA : plutôt que de programmer les règles (souvent beaucoup plus complexes qu'on ne l'aurait cru) qui président à la réalisation d'une tâche, il est désormais possible de laisser la machine les découvrir elle-même.
- Le développement de l'IA se fait également dans un contexte marqué par la "mise en données" du monde, qui touche tous les secteurs, et par le décuplement des puissances de calcul et de stockage des informations. Les applications se multiplient et touchent directement notre quotidien : reconnaissance d'images, voiture autonome, détection de maladies, recommandation de contenus, sont quelques-uns des innombrables applications qui sont en cours d'exploration.

- L'universalité de l'IA et l'infinie variété de ses déclinaisons annoncent une révolution pleine de défis et de rebonds.
- Ainsi, le sujet a largement dépassé la seule sphère scientifique. L'intelligence artificielle n'est plus un programme de recherche confiné aux laboratoires ou à une application précise. Elle est l'une des clés du pouvoir de demain dans un monde numérique.
- Voilà pourquoi il est d'intérêt général que les pouvoirs publics se saisissent collectivement de cette question. Et que la France et l'Europe puissent faire entendre leur voix. Plus que jamais, l'État doit donner un sens au développement de l'intelligence artificielle.

Une brève histoire de l'IA 1950-2008



- L'histoire de l'IA en tant que programme de recherche scientifique, est composée d'allers-retours, de phases d'emballement, d'avancées véritables et d'espoirs déçus.
- La nature des objectifs de ce programme a d'ailleurs largement évolué depuis ses débuts : elle a pu même aller jusqu'à remettre en cause l'idée d'une continuité de définition entre ce moment-là et aujourd'hui.
- Mais, la réflexion sur l'intelligence artificielle a, dès les débuts, été au cœur du développement de l'informatique. Son histoire est donc ancienne et tumultueuse.

1ère période : L'IA symbolique

- Ainsi, l'acte de naissance de l'intelligence artificielle, c'est-à-dire la première véritable utilisation de ce concept et son inscription dans un programme global de recherche est, de manière consensuelle, fixée à l'été 1956.
 - A l'université américaine de Darmouth, dans le New-Hampshire se tient alors une conférence, d'une durée de six semaines, sous l'impulsion de deux jeunes mathématiciens, Marvin Minsky et John McCarthy.
 - Elle a pour objectif de formaliser une discipline de recherche autour d'un objectif : simuler, via des systèmes de calcul, des capacités cognitives, en donnant une cohérence programmatique à différents domaines de recherche en pleine ébullition : la cybernétique (l'étude des « communications et de leurs régulations dans les systèmes naturels et

- artificiels »), le traitement complexe de l'information, les réseaux neuronaux formels (machines cherchant à imiter le fonctionnement du cerveau), la théorie des automates, les modèles de prise de décision...
- Mais ce moment inaugural s'appuie évidemment sur les recherches plus anciennes qui ont jeté les bases de la science informatique durant toute la première moitié du 20ème siècle ; la recherche concernant les fondements des mathématiques ou encore dans les années 1950 la fameuse machine de Turing qui constitue le premier modèle théorique du fonctionnement d'un ordinateur. Elle permet de penser des procédures mécanisables, suivant un ensemble de règles, pour résoudre des problèmes de calcul.
- Le développement de l'informatique a immédiatement été lié à des questionnements sur la capacité des machines à simuler des fonctions cognitives réservées jusque-là à l'être humain. En 1950, dans *Computing Machinery and Intelligence*, Turing pose en effet la question suivante : les machines pensent-elles ? Il y répond par la construction du célèbre test de Turing fondée sur la faculté d'une machine à imiter la conversation humaine.
- La première période de l'histoire de l'IA, qui est immédiatement consécutive à la conférence de 1956, est caractérisée par la prédominance de ce qu'on appelle l'IA symbolique, c'est-à-dire d'une IA manipulant des symboles logiques suivant des règles définies à l'avance.
 - Elle s'inscrit dans la droite ligne de la machine de Turing : il s'agit de créer des systèmes de démonstration, capables de résoudre des problèmes en effectuant des calculs logiques codifiables.
 - On pensait alors que l'on pourrait créer une intelligence artificielle générale à l'aide d'un algorithme universel : c'est un algorithme de ce type qu'ont créé Herbert Simon, Cliff Shaw et Allen Newell en 1959, le general problem solver.
 - Naissent ainsi dès les années 60 les premiers laboratoires de recherche en IA grâce à la DARPA au MIT et à Stanford et des chercheurs mettent au point le premier robot contrôlé par un ordinateur nommé Sharkey.

2ème période : L'IA symbolique renouvelée

- A cette première période, pleine d'excitation, a succédé un premier fort ralentissement des recherches en IA, à partir de 1973 : c'est ce que l'on a appelé le premier hiver de l'IA. Il a été dû à l'incapacité des systèmes symboliques de tenir les promesses enthousiastes qui étaient faites. Le rythme d'avancement était beaucoup plus long que prévu et surtout les systèmes étaient confrontés à un problème majeur.

- A partir des années 1980, s'ouvre une deuxième période pour la recherche en intelligence artificielle, mais qui ne change pas fondamentalement ses orientations et demeurent focalisée sur l'IA symbolique. D'autres recherches suivent en parallèle leurs cours, en développant des approches dites connexionnistes, qui se concentrent davantage sur les réseaux de neurones, mais ne constituent pas le coeur de la recherche en IA de ces années-là. Dans les années 1980, c'est sur les systèmes experts que se concentre la recherche. Ces systèmes, qui constituent davantage des promesses que des réalisations véritables, tentent de dépasser un certain nombre de limites qu'avaient rencontrées les calculateurs précédents et se focalisent notamment sur la gestion de l'incertitude et la capacité à tirer profit des connaissances acquises au cours du processus. Leur objectif est de fournir une aide à la décision et ils se distinguent des modèles antérieurs en ce qu'ils fonctionnent grâce à une base de connaissances, spécifiques à un domaine par exemple, et un moteur d'inférence générique, qui permet de déduire de nouveaux faits à partir de cette base de connaissances.
- Néanmoins le renouveau des recherches en intelligence artificielle aura été de courte durée. En 1985 s'ouvre un second hiver de l'IA. En effet ces systèmes affrontent plusieurs difficultés : leur mise en place est très complexe et très chère, tout d'abord. Et ils sont très lourds et sensibles au bruit, c'est-à-dire aux données parasites qui ne sont pas directement utiles. Il est très difficile pour ces systèmes, qui n'ont pas de fonction d'apprentissage, de discriminer entre les connaissances utiles et les connaissances inutiles, qui viennent modifier les inférences. En outre, les recherches en sciences cognitives et en philosophie de l'esprit faisaient apparaître que la conception même de l'intelligence que l'IA symbolique véhiculait était pour le moins naïve et que la réduction des opérations cognitives à des manipulations de symboles formels n'était pas la voie la plus adaptée pour simuler des comportements intelligents.
- Malgré tout, cette période est marquée par des réussites majeures comme la victoire d'un ordinateur au backgammon en 79 ou encore Mycin un système expert pour identifier des bactéries et Sachem un système d'aide à la conduite des hauts fourneaux en marche.

3ème période : NEWFAI vs GOFAI / IA connexionniste et statistique vs IA symbolique

- A partir des années 1990, s'ouvre une nouvelle période de la recherche en intelligence artificielle, qui utilise de nouvelles techniques qui se distinguent des précédentes, qui sont désormais désignées sous le terme de GOFAI : *good-old fashioned artificial intelligence*. Ces techniques sont diverses mais relèvent toutes de l'apprentissage. Elles tentent donc de modéliser les processus cognitifs non à partir d'un ensemble de règles formelles logiques, mais à partir de boucles

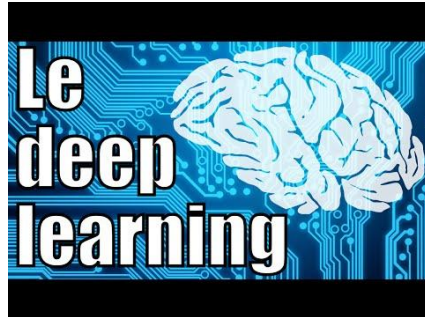
- d'apprentissage et retour d'information du monde extérieur qui semblent davantage correspondre à la manière dont le développement de l'intelligence humaine s'effectue. Un cas extrême est celui où la boucle est réduite à sa plus simple expression : le système apprend d'un seul coup à partir d'exemples du phénomène observé (apprentissage supervisé). L'une de ces approches, appelée connexionniste, existait depuis les années 1950 avec l'outil de classification statistique Perceptron, mais avait elle aussi été jugée sans avenir dans ses versions initiales (linéaires).
- L'invention de l'algorithme d'apprentissage dit de *rétro-propagation du gradient* a vu un renouveau au début des années 90.
- En parallèle, Vladimir Vapnik posait les bases théoriques des Support Vector Machines (SVM), et obtenait également des succès expérimentaux indéniables face aux réseaux de neurones de l'époque. Ce fut l'âge d'or des SVM, des années 1990 aux années 2000.
- Par ailleurs, il était clair (même si sans justification théorique) que les réseaux de neurones devaient être capables de traiter des problèmes de grande dimension (tant en nombre de variables qu'en taille des données disponibles) en augmentant le nombre de couches. C'est ainsi que grâce à la ténacité de quelques visionnaires (Geoff Hinton, Joshua Bengio et Yann LeCun), et à l'augmentation des capacités de calcul, des variantes de la rétro-propagation du gradient ont pu être mises au point et rendues efficaces (même s'il faut des centaines de GPU pendant des centaines d'heures pour apprendre des grands réseaux sur des grandes bases d'exemples), et les réseaux obtenus, baptisés *Réseaux Profonds* ont commencé à surpasser toutes les autres approches, en particulier dans le domaine de la vision par ordinateur et de l'analyse d'image, puis dans le domaine du langage naturel (en particulier la traduction automatique). L'apprentissage profond tient aujourd'hui le haut du pavé pour l'apprentissage supervisé de très grande taille.
- Il faut aussi parler des autres branches de l'IA comme le Web sémantique qui ne sont pas morts avec les systèmes experts
- Les années 90 et 2000 ont elles aussi été marquées par de belles réussites : les débuts des véhicules autonomes en 94-95, la victoire du programme Deep Blue aux échecs en 97, Asimo le premier robot humanoïde développé par Honda capable de marcher ou encore le développement de l'IA Watson par IBM dès les années 2005.

Depuis 2008 : l'explosion du *deep learning*

mai 2017 :
Victoire d'AlphaGo
contre Ke Jie
décembre 2017:
Alpha Zero



- Depuis 2008 trois facteurs combinés ont été à l'origine d'avancées extrêmement considérables dans le domaine de l'intelligence artificielle : l'amélioration des algorithmes d'apprentissage sur les réseaux de neurones, l'explosion du nombre de données disponibles et des capacités de calcul.
- Une illustration : **AlphaGo** est un programme informatique capable de jouer au jeu de go, développé par l'entreprise britannique Google DeepMind. Le jeu de go offre 130 000 déplacements possibles après le premier coup et compte la bagatelle de 10 puissance 170 positions.
- Il s'agit d'une étape symboliquement forte puisque le programme joueur de go est alors un défi complexe de l'intelligence artificielle. En mars 2016, il bat Lee Sedol, un des meilleurs joueurs mondiaux (9^e dan professionnel). Le 27 mai 2017, il bat le champion du monde Ke Jie et sa retraite est annoncée.
- L'algorithme d'AlphaGo combine des techniques d'apprentissage automatique et de parcours de graphe, associées à de nombreux entraînements avec des humains, d'autres ordinateurs, et surtout lui-même.
- Cet algorithme sera encore amélioré dans les versions suivantes. AlphaGo Zero en octobre 2017 atteint un niveau supérieur en jouant uniquement contre lui-même. AlphaZero en décembre 2017 surpasse largement, toujours par auto-apprentissage, le niveau de tous les joueurs humains et logiciels, non seulement au go, mais aussi aux échecs et au shōgi.



Pour voir la vidéo, cliquer sur l'image



- Face aux récentes évolutions de l'IA, les Français sont réalistes car 93% d'entre eux considèrent que les technologies de l'IA ont ou auront un impact important sur nos vies au quotidien.
- Pour près de la moitié d'entre eux, c'est même une révolution.

DÉMARCHES & SERVICES

69%



L'I.A peut aider aux
**démarches
administratives**

Pour quels services ?

- ✓ Renouvellement automatique des documents officiels
- ✓ Remplissage automatique des autres déclarations administratives
- ✓ Remplissage automatique des déclarations fiscales

33%



Qui au
**partage de données
personnelles**

- ✓ Auprès d'entreprises ou d'institutions identifiées
- ✓ Intérêt général
- ✓ Nouveaux services personnalisés

I.A* = Intelligence Artificielle

**Mission
Villani**

Données recueillies et analysées par Médiamétrie.
Étude réalisée du 1^{er} au 7 février et du 15 au 19 mars 2018 via Touch By Médiamétrie.
Échantillon de 3 126 internautes âgés de 15 ans et +, représentatif de la population d'internautes vivant en France.
Tous droits réservés © 2018 Médiamétrie - www.mediametrie.fr


Médiamétrie

IA FAIBLE

C'est la capacité à reproduire des tâches humaines spécifiques et délimitées.

Cela rend nos vies **plus faciles** !

IA FORTE

C'est la capacité d'égaliser/dépasser des tâches humaines dans tous les domaines simultanément.

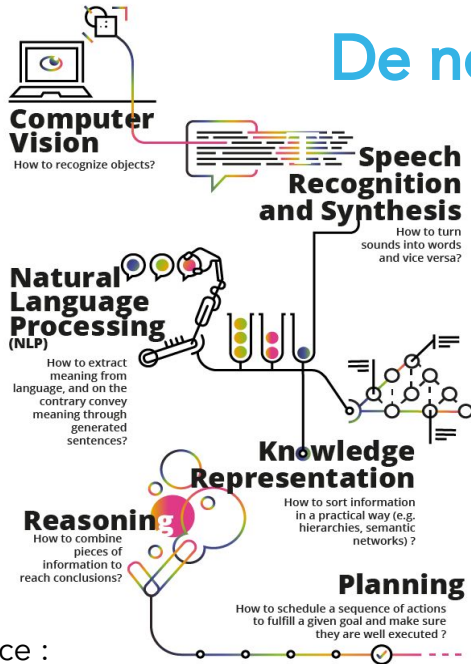
Nous en sommes **loin** !

Quel est l'objectif de la recherche en intelligence artificielle ?

- Après tout pourquoi créer des machines intelligentes ? A minima, on souhaite qu'elles rendent nos vies plus faciles, en prenant en charge les tâches complexes et/ou répétitives du quotidien : trouver le meilleur itinéraire pour se rendre chez des amis, sélectionner les contenus (musique, chanson, articles) conformes à nos goûts, traduire des textes etc... Dans le jargon des experts, **c'est ce qu'on appelle l'intelligence artificielle faible**.
- On la qualifie de "faible" car ces programmes ne fonctionnent que sur des tâches spécifiques et délimitées. Concrètement, les algorithmes qui créent nos *playlists* favorites, sélectionnent les meilleurs résultats de recherche ou taguent nos amis sur les médias sociaux ne savent faire que cette tâche ! Ça n'enlève rien à leur mérite, même si cela peut conduire à une forme de bêtise artificielle : si une technique d'IA est conçue pour dire si une image contient un chien ou un chat, lui soumettre une photo d'un humain la conduira à donner la réponse "chien" ou "chat" malgré l'énormité apparente de l'erreur.
- **A l'opposé une IA forte aurait la capacité d'égaliser ou dépasser les capacités humaines dans tous les domaines simultanément.** Nous en sommes loin : il suffit de quelques minutes de conversation avec les agents conversationnels (chatbots) les plus récents pour s'en convaincre. Néanmoins on observe déjà des vives polémiques dans le débat public autour du potentiel destructeur ou positif d'une

- telle IA. Ces discussions puisent dans un imaginaire largement infusé par la science fiction, qui nourrit le développement de l'IA depuis longtemps. Les innovations technologiques antérieures ont souvent alimenté ce type de crainte et il n'est pas étonnant qu'il en soit de même avec le développement de l'IA. Il n'en reste pas moins que les pouvoirs publics sont confrontés à cette question de manière pressante, du fait d'un certain nombre de projections, par exemple sur l'emploi. Réguler un objet si mouvant et fantasmatique n'est pas une mince affaire.

De nombreuses applications

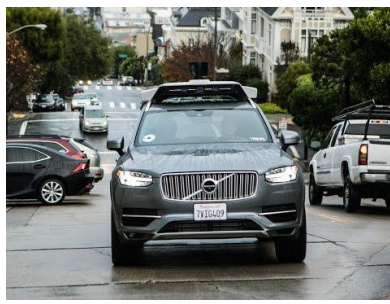
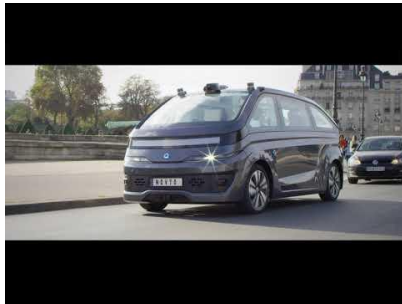


Source :
Fabernovel

L'apprentissage profond

- L'apprentissage profond (*deep learning* en anglais) est un sous-domaine de l'apprentissage automatique qui a connu un véritable renouveau depuis 2006. Ces modèles sont construits à partir d'éléments de calcul de base appelés neurones, disposés en général en couches successives. Ces neurones sont reliés entre eux par des connexions dont l'importance est ajustée durant la phase d'apprentissage.
- Chaque neurone transforme ses entrées en une sortie en appliquant une fonction de transfert à l'ensemble des informations qui lui parviennent par ces connexions. Les réseaux de neurones sont des systèmes schématiquement inspirés du fonctionnement du cerveau. Toutefois, il ne s'agit que d'une analogie grossière.
- Après apprentissage, on constate que le réseau de neurones a décomposé les données d'entrée en différents couches d'abstraction hiérarchisées. Par exemple, pour identifier des visages sur une photo, la première couche identifie uniquement les formes élémentaires constituant les visages (lignes, cercles, etc..), la couche suivante les assemble en gros traits du visage (nez, lèvres, front etc..) et la description est progressivement affinée (taille du nez, largeur des narines etc..) jusqu'à recomposer complètement les visages.

Les réseaux de neurones sont particulièrement efficaces dans le traitement de données assimilables à des signaux : son, image, langage, vidéos car ils sont capable d'en capturer les structures complexes sur les plan spatial et temporel.



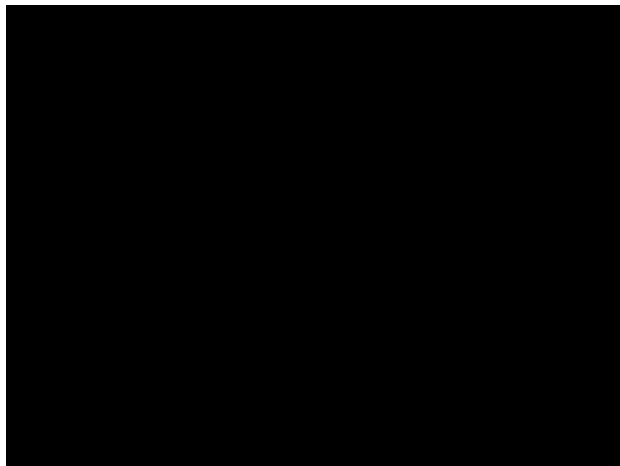
Pour voir les vidéos, cliquer sur les images

Détection du cancer



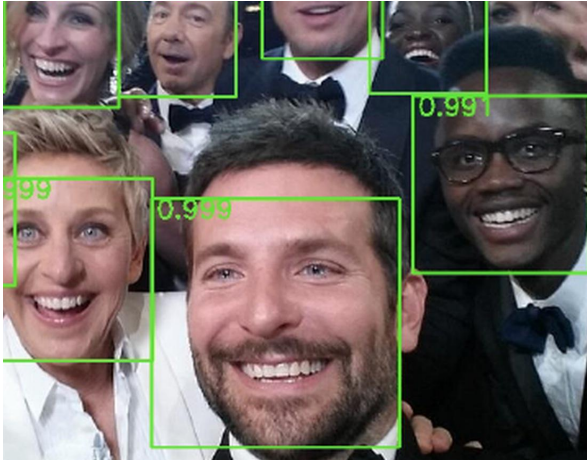
Pour voir les vidéos, cliquer sur les images

La reconnaissance vocale smartphones peut
écrire des messages trois fois plus vite que l'humain

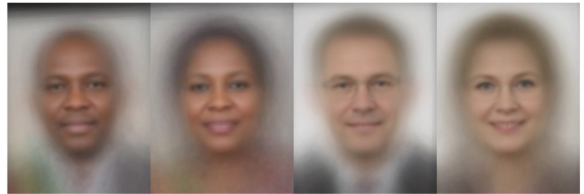


Pour voir la vidéo, cliquer sur l'image

Reconnaissance d'images



Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%



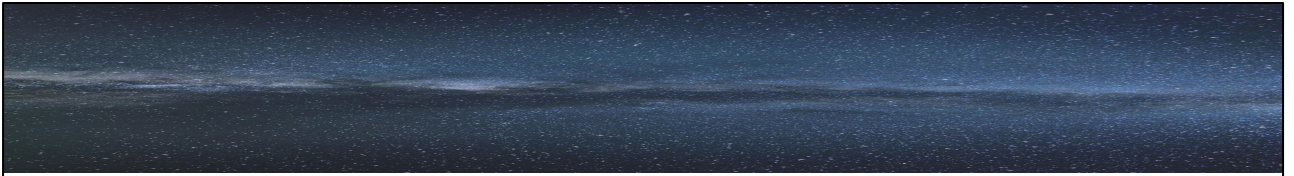
Reconnaissance d'images

- Interpréter une image — reconnaître une personne ou un objet et l'environnement avoisinant — est une chose relativement facile pour un humain. Chaque jour, notre cerveau traite sans grand effort des informations visuelles complexes : une photo de famille, une voiture, un paysage. Mais pour un ordinateur, c'est une tâche très difficile.
- Pourtant l'enjeu est de taille, car le développement de la voiture autonome (pour la perception de son environnement), l'annotation automatique d'images, l'amélioration des systèmes d'identification, la détection de pathologies à partir de dispositifs d'imagerie médicale sont tous dépendants des progrès en matière de reconnaissance d'images et de vidéos.

Comment les médias sociaux reconnaissent-ils les visages des personnes présentes sur une photo?

- Ils ont recours à l'apprentissage automatique. Celle-ci **comporte deux phases**: la première est celle de l'apprentissage à proprement parler et consiste à choisir un modèle, **ici réseaux de neurones profonds convolutionnels**, puis ajuster ses paramètres à partir de données en entrée - par exemple les visages des amis de l'utilisateur pour lesquels il dispose d'images étiquetées.. La deuxième phase est celle de l'inférence. A partir des paramètres qui ont été appris, l'algorithme effectue

- la tâche qui lui a été fixée - par exemple taguer les visages présents sur une photo nouvellement téléchargée sur la plateforme.
- Cela étant la reconnaissance d'image n'est pas exempte de biais et une chercheuse du MIT media Lab a démontré que les algorithmes de reconnaissance faciale étaient plus performants sur des hommes que sur des femmes et sur des peaux claires que foncées.



Pour voir les vidéos, cliquer sur les images

<https://www.youtube.com/watch?v=ohmajJTcpNk&feature=youtu.be>

Welcome to the beta version of Lyrebird

Lyrebird allows you to create a digital voice that sounds like you with only one minute of audio.

Create your digital voice

Or log in if you already have an account

This **beta version** allows anyone to create their digital voice with **only one minute** of audio. Simply sign up, record yourself for at least one minute and you will be able to generate any sentence you like with your digital voice.

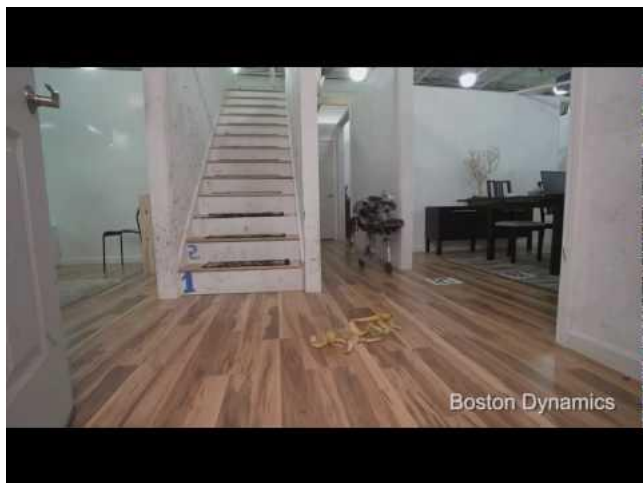


THIS IS AN ARTIFICIAL VIDEO CREATED BY LYREBIRD.
IT DOES NOT CONVEY THE OPINION OF BARACK OBAMA.

Pour voir la vidéo, cliquer sur l'image

<https://lyrebird.ai/>

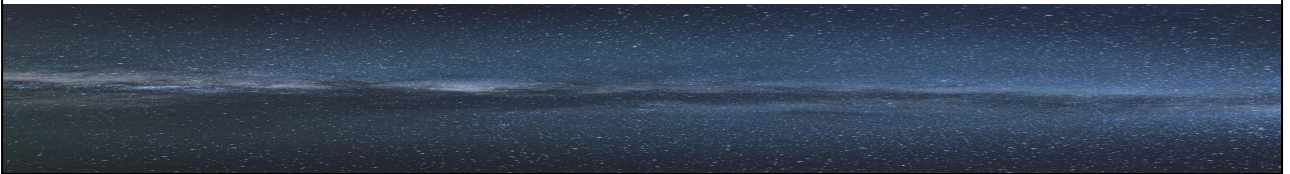
BONUS : L'arme fatale contre l'IA, la peau de banane



Pour voir la vidéo, cliquer sur l'image

Donner un sens à l'IA

La vision de la mission



- Plus que jamais, l'État doit donner un sens au développement de l'intelligence artificielle. **Donner un sens, c'est-à-dire donner un cap, une signification et des explications.** Voilà l'objectif de ce rapport.
- **Donner un sens, c'est tout d'abord donner un cap.** C'est l'objectif de la structuration proposée pour la politique industrielle : quatre secteurs prioritaires sont définis. Donner un cap c'est garder la place décisive de la France dans la recherche en IA. C'est aussi veiller à ce que l'IA mette en son coeur la préoccupation écologique.
- **Donner un sens, c'est également donner une signification.** L'intelligence artificielle doit prendre en compte plusieurs aspects.
 - Tout d'abord la nécessité de penser les modes de complémentarité entre l'humain et les systèmes intelligents. Au cœur du développement de l'IA doit résider la nécessité de mettre en œuvre une complémentarité qui soit capacitante, en ce qu'elle permet de désautomatiser les tâches humaines. Pour favoriser la transition des tâches et des métiers dans ce sens, des expérimentations devront être mises en place sur l'ensemble des territoires, notamment à destination des populations les plus touchées par l'automatisation.
 - Dans un monde marqué par les inégalités, l'intelligence artificielle ne doit pas conduire à renforcer les phénomènes d'exclusion et la concentration de la valeur. L'IA doit prioritairement nous aider à activer nos droits fondamentaux, augmenter le lien social et renforcer les solidarités. La mixité doit être également au cœur des priorités : la situation est alarmante dans les filières numériques, tant les femmes

- sont peu représentées. Les algorithmes peuvent en outre reproduire des biais sexistes.
- Enfin une société algorithmique ne doit pas être une société de boîtes noires : l'intelligence artificielle va être amenée à jouer un rôle essentiel dans des domaines aussi variés que cruciaux (santé, banque, logement,...) et le risque de reproduire des discriminations existantes ou d'en produire de nouvelles est important. À ce risque s'en ajoute un autre : la normalisation diffuse des comportements que pourrait introduire le développement généralisé d'algorithmes d'intelligence artificielle. Il doit être possible d'ouvrir les boîtes noires, mais également de réfléchir en amont aux enjeux éthiques que les algorithmes d'intelligence artificielle peuvent soulever.
- **Donner un sens c'est enfin expliquer** : expliquer ces technologies à l'opinion pour la démystifier — le rôle des médias est de ce point de vue primordial —, mais aussi expliquer l'intelligence artificielle en elle-même en développant les recherches sur l'explicabilité.

6 AXES

- Une politique économique ambitieuse basée sur les données
- Pour une recherche agile et diffusante
- Anticiper et maîtriser les impacts sur le travail et l'emploi
- Une IA écologique au service d'une économie durable
- Quelle éthique pour l'IA ?
- Pour une IA inclusive et diverse

5 FOCUS SECTORIELS

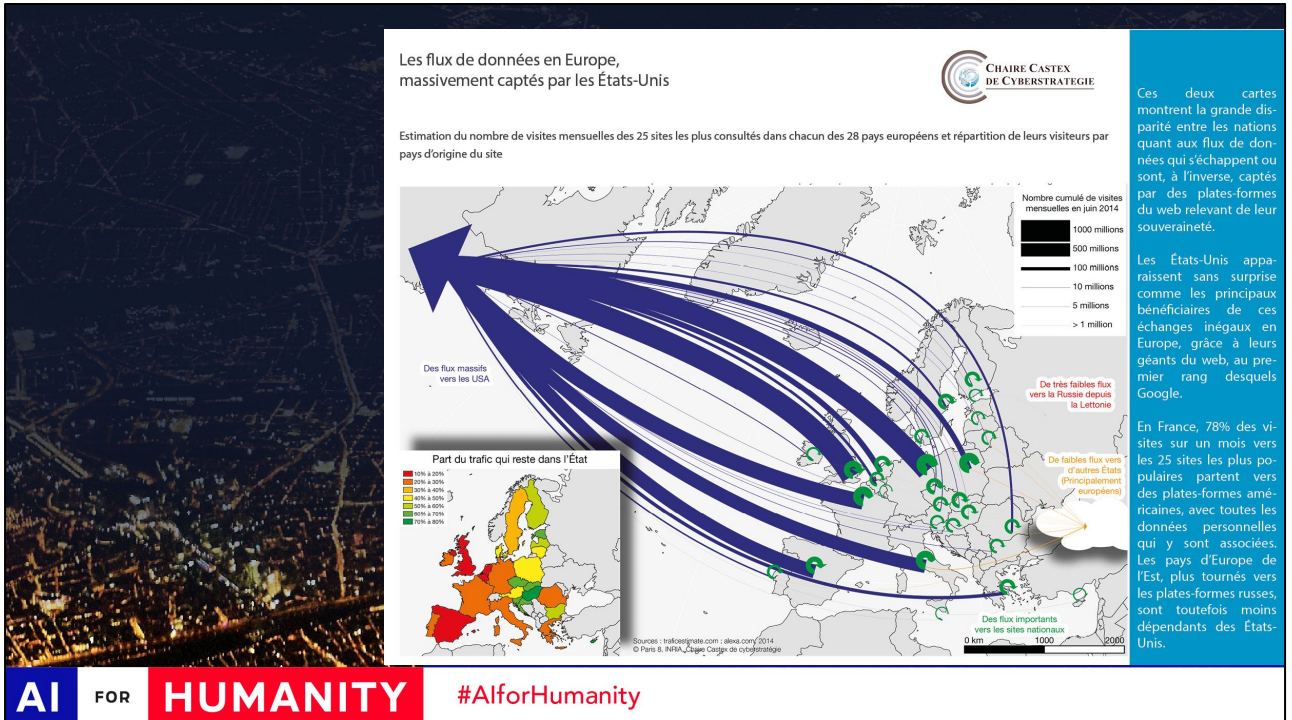
- L'Education
- La santé
- L'agriculture
- Transport
- Défense et sécurité

- Le rapport comprend 6 axes : économie et données, recherche emploi et travail, écologie, éthique et mixité et inclusion ainsi que 5 focus sur des secteurs clefs que nous allons vous présenter.



1

Faire émerger un
écosystème français et
européen de l'IA appuyée
sur la donnée



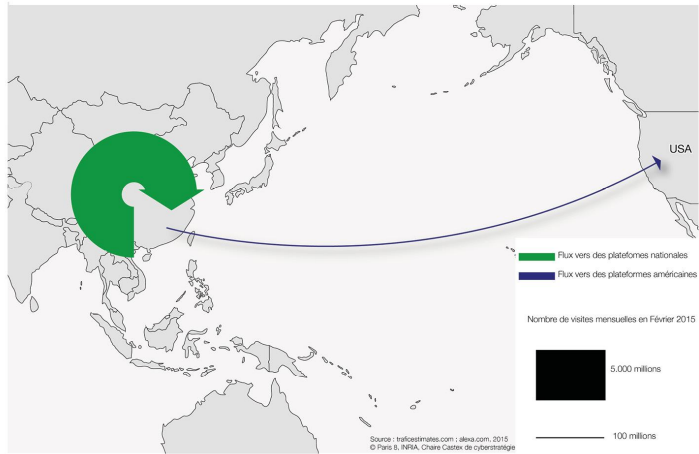
- Tout d'abord un chiffre clef qui démontre de la plateformeisation de notre économie et du règne des GAFAM : 80% des sites visités par les français et donc de leurs données sont captées par des entreprises américaines.

Donnée

Les flux de données en Chine restent principalement en Chine



Estimation du nombre de visites mensuelles des 25 sites les plus consultés en Chine et répartition de leurs visiteurs par pays d'origine du site



La Chine, de son côté, parvient à capter l'essentiel des données de ses utilisateurs.

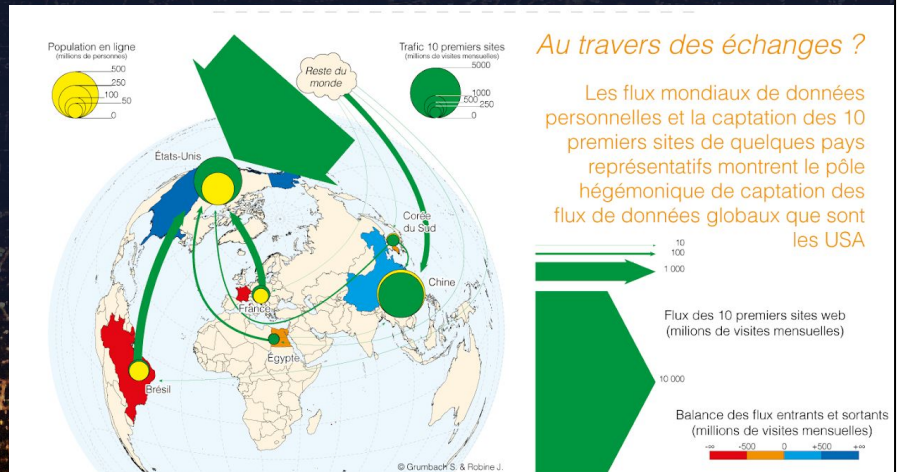
Cela résulte de la politique volontariste du gouvernement chinois de favoriser l'émergence de ses propres champions du numérique dont Baidu, Alibaba et Weibo.

Ces cartes sont toutefois expérimentales. Mesurer et cartographier l'ensemble des flux de données requiert des métriques qui n'existent que partiellement aujourd'hui.

L'enjeu est de comprendre comment les géants du numérique influencent le paysage géopolitique et comment ils transforment l'exercice de la souveraineté des Etats.

AI FOR HUMANITY

Donnée



AI FOR HUMANITY

#AIforHumanity

Toute solution d'intelligence artificielle repose sur trois piliers :

- Un algorithme
- Des moyens de calculs
- Des données

Diagnostic: Toute solution d'intelligence artificielle repose sur trois piliers :

1. Un algorithme

Les algorithmes d'apprentissage automatique sont souvent disponibles dans des grandes bibliothèques (TensorFlow et Theano) accessibles en open source. Ces outils sont gratuits, bien documentés et soutenus par des communautés dynamiques. La disponibilité de ces outils a rendu les applications d'apprentissage automatique des machines beaucoup plus accessibles aux développeurs qui ne travaillent pas pour les géants de la Tech.

2. Des moyens de calculs

La deuxième piliers concerne les moyens de calcul, à la fois sous la forme de la puissance des processeurs disponibles et des solutions de stockage de données. En matière d'apprentissage automatique, la vitesse et la performance de la phase d'apprentissage sont conditionnées par l'ampleur des moyens matériels qui leur sont alloués, notamment en matière de processeurs dédiés (aujourd'hui principalement des GPU). La dimension des infrastructures conditionne donc la productivité et l'efficacité de la recherche et du développement. La seconde phase est celle de l'inférence, beaucoup plus légère en besoin de ressources matérielles. Les géants du numérique disposent d'important moyens de calculs qu'ils mettent à disposition aux entreprises moyennant finance. Toutefois, pour éviter toute dépendance au GAFAM dans ce domaine, il est essentiel d'investir massivement dans des supercalculateurs. Nous faisons des recommandations en ce sens dans le rapport.

Mais avant de vous mettre en quête du bon algorithme et d'investir dans des moyens de calculs importants, vous avez besoin de données, et préférablement en grande quantité car la qualité et la masse des données détermineront le niveau d'applications IA que vous pouvez développer. Or, c'est là que le bât blesse. Aujourd'hui **on constate une asymétrie colossale des flux de données entre l'UE et les états Unis**

- Rien qu'en France, près de 80% des visites vers les 25 sites les plus populaires sur un mois sont captés par les grandes plateformes américaines.
- Situation unique au monde car la Russie et la Chine ont des politiques volontaristes de captation des données de leurs utilisateurs.

Pour réduire cette asymétrie, **il est nécessaire de développer une politique européenne souveraine de la donnée, conforme à notre cadre légal (RGPD) et nos valeurs.**

Que doit faire la puissance publique pour faciliter ce mouvement ?

- Accompagner les acteurs de l'écosystème
- Intervenir au niveau réglementaire : DIG, TDM, Portabilité

On préconise deux types d'intervention:

- La première doit prendre la forme d'un accompagnement des acteurs à travers trois leviers d'action :
 1. **La promouvoir d'un modèle de collaboration et de partage entre acteurs d'un même secteur.**
 - **Benchmark** : l'entreprise Dawex accompagne les entreprises dans la contractualisation de leurs échanges en données (contrat de licence, durée, territoire, usages, capacité à sous-licencier, etc.) et s'assure qu'ils respectent les réglementations en vigueur (notamment le RGPD).
 2. **L'organisation d'événements sectoriels pour sensibiliser et inciter au partage et à la mise en commun des données.**
 - **Exemple d'évènement** : « Forum de la data », lieu de dialogue, d'idéation et pour certains d'acculturation entre acteurs souhaitant partager leurs données.
 3. **L'accompagnement et le conseil des entreprises dans la démarche de contractualisation de leurs échanges en données**
- **La seconde forme d'intervention porte sur le volet réglementaire.** L'Etat doit:
 1. **organiser l'ouverture au cas par cas de données détenues par des acteurs privés, si elles sont d'intérêt général**
 - Pour cela, elle peut s'appuyer sur les dispositions légales déjà existantes:
 - Dont la notion de données d'intérêt général (DIG)

- contenue dans la loi pour une République numérique s'appliquant pour:
 - les entreprises concessionnaires de service public
 - les entreprises gestionnaires des réseaux publics de distribution de gaz naturel ou d'électricité ou à des fins statistiques.
- Et certaines données susceptibles de faciliter la transition énergétique conformément à la loi Macron.

Warning : la création d'un éventuelle d'un statut général de « données d'intérêt général » n'apparaît pas souhaitable à ce stade car il constituerait une atteinte injustifiée à la liberté d'entreprendre. Toutefois, la plupart des acteurs auditionnés par la présente mission se montrent favorables à une ouverture progressive, au cas par cas et selon les secteurs, de certains jeux de données pour des motifs d'intérêt général.

1. Favoriser les pratiques de fouille de texte de données

- **Définition:** la **fouille de texte de données** désigne l'ensemble de traitements informatiques consistant à extraire des connaissances selon un critère de nouveauté ou de similarité dans des textes ou des bases de données.
 - **Quel est l'enjeu ?** ces techniques ont des potentiels énormes pour la découverte scientifique et le développement de nouvelles connaissances.
 - **Quels sont les barrières?** Aujourd'hui, la reproduction des bases de données nécessaire à la mise en place de systèmes de fouille suppose un accord explicite du propriétaire de l'œuvre ou du détenteur des bases de données concernées, y compris lorsque l'accès à ces données est licite (chercheurs accédant à une base de données via ses identifiants universitaires).
- **Que faire?** Développer un cadre légal permettant la fouille de texte de données pour les chercheurs, journalistes et éventuellement les entreprises.

2. Mettre en oeuvre la portabilité de toutes les données liées à un compte utilisateurs dans une visée citoyenne, conformément à l'article 48 de la Loi pour une République numérique.

- **Ici, les objectifs sont de :**
 - Constituer de nouvelles bases de données à usage de service public ;
 - Contribuer à donner un sens nouveau au droit à la portabilité, en permettant une meilleure circulation des données sous le contrôle exclusif des citoyens

Il est nécessaire de **réfléchir activement à la réforme du cadre international applicable aux transferts de données** car le privacy shield ne donne pas suffisamment de garanties à

la protection des données personnelles des Européens.

- 1 Encourager la **mutualisation des données** par secteurs
- 2 Consolider et rendre visible **l'écosystème français et européen de l'IA**
- 3 Faire levier sur **la commande publique**

- **La mutualisation des données demeure encore une démarche exceptionnelle**
 - Selon une étude menée par le think tank European Digital Forum en 2017, 90% des entreprises interrogées affirmaient ne pas partager leurs données avec d'autres
 - Mais ce mouvement d'ouverture est en marche, comment en témoigne diverses initiatives lancées par des entreprises et de fondations publiques : Uber Movement, AirBnB, DataVille, OpenStreetMap.
- **Il faut soutenir et encourager ce mouvement, particulièrement dans les secteurs où les entreprises européennes sont leaders pour :**
 - **des raisons économiques:**
 - Le premier acte de la « bataille de l'IA » portait sur les données à caractère personnel (bataille gagnée par les grandes plateformes)
 - Le second acte va porter sur les données sectorielles (données de santé, données éducatives, données de transport) : c'est sur celles-ci que la France et l'Europe peuvent se différencier.
 - **Mais également des soucis de sécurité:**
 - Prenons l'exemple de la voiture autonome car il est particulièrement parlant. Aujourd'hui, les constructeurs automobiles développent chacun leurs propres modèles d'apprentissage. Ce qui n'est pas optimal car ça les conduit à développer des référentiels adaptés uniquement à leurs véhicules. Or, l'intérêt de tout citoyen est de bénéficier du niveau maximal de sécurité, indépendamment du modèle choisi. La sécurité étant elle-même dépendante de la robustesse du référentiel développé, les

- constructeurs devraient plutôt mutualiser leurs données pour augmenter le niveau de sécurité de tous les véhicules.
- Les politiques de mutualisation des données doivent s'accompagner d'un véritable effort de consolidation de l'écosystème français et européens de l'IA. Pour y parvenir, **nous suggérons différentes mesures** :
 1. Créer un guichet unique pour accompagner les futurs acheteurs de solutions d'IA
 2. Mettre en place un label rattaché à la French Tech et des prix d'innovation de terrain : ils devraient désigner des acteurs français de l'IA.
 3. Consolider les filières d'acheteurs et limiter autant que possible les achats extraeuropéens quand il existe de meilleures alternatives.
 4. Faciliter le dialogue entre les acteurs de l'IA et les régulateurs (ARCEP, CNIL, ANSSI)
 - **Benchmark: French Tech Central** (lieu de passage, d'information et de rencontre pour toutes les startups françaises, situé dans le campus de startups Station F. 30 services publics y sont présents en permanence ou à temps partiel, avec pour mission de conseiller et orienter les entreprises.
 - 5. Inciter à la constitution d'écosystèmes centrés sur les filières utilisatrices d'IA, animées par de grands groupes et intégrateurs du domaine au travers de partenariats d'entreprises.
 - **Dans ce domaine, trois leviers sont activables:** organisation du dialogue des filières utilisatrices d'IA, incitation dans le cadre de la commande publique à la constitution de groupements d'entreprises ; la création d'un dispositif d'information pour aider les entreprises à trouver des partenaires dans le cadre de réponses à des appels d'offres publics ou privés.
 - De la même façon que les acteurs privés doivent se saisir des problématiques de l'IA et en devenir des acheteurs, la puissance publique devra également y avoir recours pour ses propres besoins. Et quand on considère les la masse des sommes dépensées par la commande publique (ex: 71,5 milliards d'euros annuels pour l'État, les établissements publics et les collectivités), l'enjeu apparaît considérable.
 - **Ainsi, il nous semble nécessaire d'encourager l'achat public de l'IA pour remplir trois objectifs:**
 - Satisfaire certains besoins de l'État en matière d'IA
 - Soutenir l'écosystème avec la commande publique
 - L' aider à constituer une vitrine exportable en Europe et à l'international.

Nous avons formulé 5 recommandations pour faciliter la commande publique:

1. **Ramener les seuils d'application de l'ordonnance sur les marchés publics aux niveaux européens**
 - a. **Seuils européens actuels:** 144 000 euros HT pour les marchés publics de fournitures et de services de l'état ; 221 000 euros HT pour les marchés publics de fournitures et de services des collectivités territoriales et pour les marchés publics de fournitures des autorités publiques centrales opérant dans le domaine de la défense ; 443 000 euros HT pour les marchés publics de fournitures et de services des entités adjudicatrices)
 - b. **Seuil national** : 25 000 HT
2. **Mettre l'achat public au service du soutien des industries européennes**
 - **Pour cela nous suggérons** d'introduire au niveau européen une possibilité pour la puissance publique, dans le cadre de la passation

- technologique européenne pour, par exemple, privilégier un acteur européen lorsque le déséquilibre de la concurrence est manifeste.
1. **Il est essentiel de dynamiser l'achat public innovant.**
 - Deux options possibles:
 1. **Partenariat d'innovation.** Dans une procédure de marchés, elle permet de couvrir le besoin de la phase de recherche amont et d'expérimentation jusqu'à la phase d'achat du produit opérationnel, sans avoir à remettre les acteurs en concurrence entre ces différentes phases.
 2. **Le dialogue compétitif.** Il s'agit d'une solution adaptée à la conclusion de marchés complexes, pour lesquels l'acheteur public ne peut définir seul et à l'avance les moyens techniques qui vont répondre à son besoin, ou encore pour lesquels il n'est pas en mesure d'établir un montage juridique ou financier adapté.
 2. **Mettre en place une protection de l'acheteur public pour l'inciter à l'ingénierie contractuelle.**
 - **Scénario:** identification spécifique de type acheteur innovant officialisant la prise de risque demandée afin qu'un échec ne soit pas pénalisant.
 3. **Généraliser l'emploi des exceptions à l'ordonnance sur les marchés publics actuellement présentes** pour les marchés de recherche et de développement, les marchés concernant les intérêts essentiels de l'État et les marchés de défense ou de sécurité



Pour renforcer l'écosystème français et européen de l'intelligence artificielle, il nous faut tirer parti des avantages comparatifs et des secteurs d'excellence de notre économie.

- Comment avons nous identifié ces secteurs stratégiques? A l'aide des critères suivants :
 - **Impact** : celui-ci doit être porteur de profondes transformations d'un point de vue économique, mais également en termes d'intérêt général ;
 - **Écosystème** : la capacité à amorcer et entretenir une dynamique impose de disposer au préalable d'un socle d'acteurs publics et privés solides sur lequel s'appuyer ;
 - **« Carburant initial »** : données, cas d'usage, connaissances métier, ressources, cadre souple, marché... Les données sont évidemment un élément essentiel et constituent un avantage comparatif important.
 - **Finances et ressources** : les secteurs identifiés doivent être en mesure de mobiliser à la fois des financements publics et des financements privés, ainsi que les ressources humaines nécessaires ;
 - **Marchés et ouverture** : les acteurs présents dans ces marchés doivent être en mesure savoir-faire sur des marchés publics, privés, en France et à l'international;
 - **Dualité et percolation des domaines** : les technologies qu'ils développent doivent pouvoir facilement être transposées dans d'autres domaines
 - **Impulsion de l'état** : il faut que les secteurs nécessitent une intervention initiale forte de l'État pour se transformer, ce qui n'est pas valable pour une grande majorité de secteurs industriels.
- En considérant ces exigences, **notre mission recommande de retenir quatre**

- **secteurs en particulier : santé, transports-mobilités, environnement et défense-sécurité.**

Dans les secteurs prioritairement retenus, cette transformation doit s'articuler sur trois axes:

1. Un soutien à l'innovation pensé de bout en bout.
2. Une mobilisation et une structuration des écosystèmes autour de grands enjeux et défis sectoriels
3. Les organisations doivent se rendre accueillantes à l'innovation, qu'elle soit technologique, d'usage ou de modèle d'affaires.
 - **Sur ce point, on peut particulièrement s'inspirer de l'exemple de la DARPA (benchmark)**

The background is a chalkboard filled with various mathematical expressions and diagrams. At the top left, there is a formula: $A_{ij} = \frac{1}{2} \mathbf{u}_i^T \mathbf{V}_j, \mathbf{u}_i^T \cdot \mathbf{0} \dots$. Below it, a summation formula: $\min_{u \in P(\Omega)} \sum_{i=1}^N \lambda_i W_p^p(u, v_i)$. To the right, a formula for a conditional probability: $\hat{\theta}_{i,j,k}^{MAP} = \hat{P}(X_i = x_k | p_a(X_i) = x_j) =$. Further right, another formula: $A_{ij} = \frac{1}{2} \mathbf{u}_i^T \mathbf{V}_j, \mathbf{u}_i^T \cdot \mathbf{0} \dots$. In the center, there is a diagram of a neural network with four input nodes labeled h_{rot} and several hidden nodes. Below the neural network, there is a formula: $E(W) = \frac{1}{P} \sum_{p=1}^P (y D P(Z^p, W) + \log)$. At the bottom right, there is a formula: $MMD_k(x, \hat{x}) = \frac{1}{n} \sum_{i,j} k_{ij}$. On the left side, there is a large white number '2' and a blue triangle. The text 'Pour une recherche agile et diffusante sur l'IA' is written in white on the right side of the number '2'.

2 Pour une recherche agile et diffusante sur l'IA

La recherche en IA fait
l'objet d'une **intense**
compétition internationale,
notamment entre **la Chine**
et les États-Unis



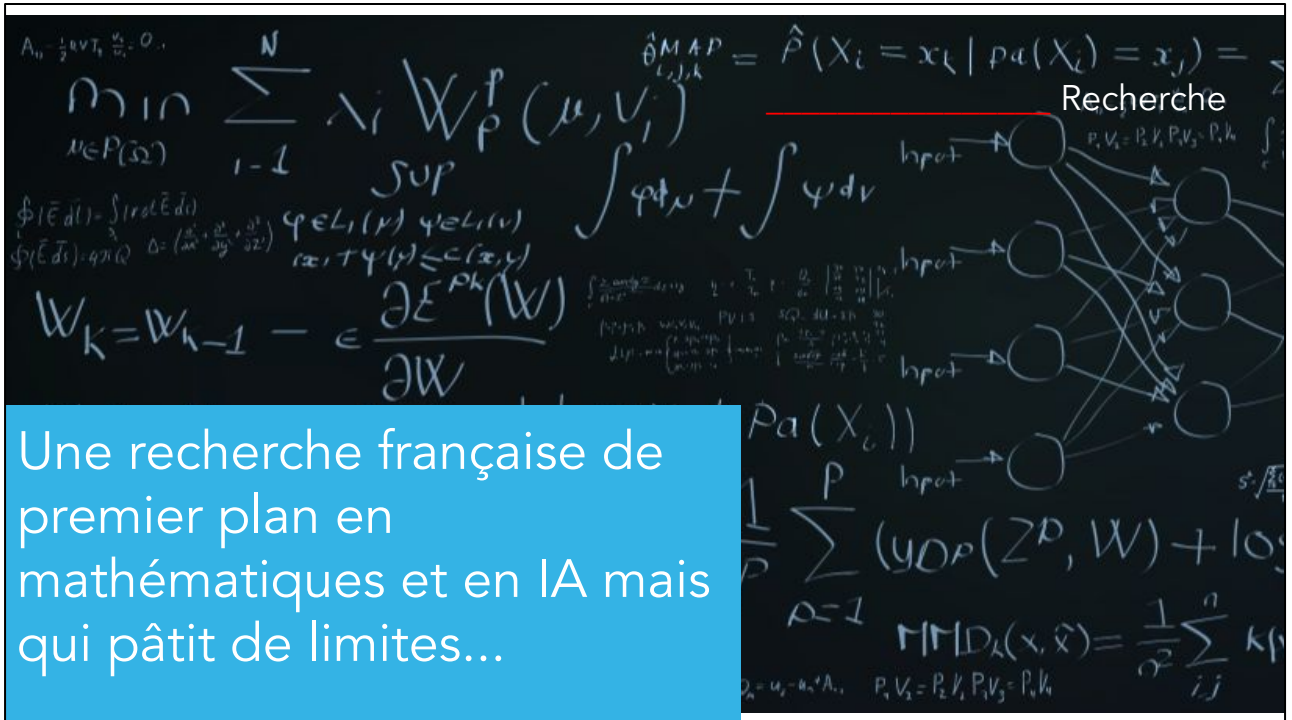
#AIforHumanity

AI

FOR

HUMANITY

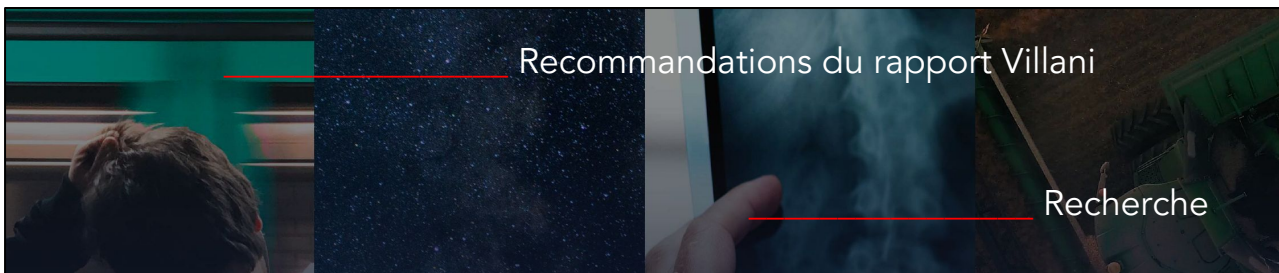
- La recherche en IA fait l'objet d'une indéniable compétition entre les pays du monde entier et plus particulièrement entre la Chine et les États-Unis.
 - Par exemple le Pentagone avait formulé une demande de financement hors DARPA de 12-15 Mds USD pour la recherche militaire en IA pour l'année 2017 ;
 - La Chine a mis en place le plan Next Generation Artificial Intelligence Development Plan en juillet 2017 visant à faire passer le poids économique de l'IA de plus de 150 milliards de yuans (22,15 milliards de dollars) à l'horizon 2020 à 400 milliards de yuans (59,07 milliards de dollars) d'ici 2025.



Une recherche française de premier plan en mathématiques et en IA mais qui pâtit de limites...

- L'ESR français en IA a toujours occupé une place de choix au niveau international. La formation scientifique française jouit d'une image d'excellence reconnue et alimente un vivier sans cesse renouvelé de chercheurs au meilleur niveau mondial.
 - Par exemple, dans les années 80, le ministère japonais de l'Économie, le MITI a fait le choix du langage de programmation français Prolog. L'essor des réseaux de neurones profonds doit beaucoup à la recherche française : le plus célèbre chercheur étant Yann Le Cun, ce dernier travaille aux États-Unis depuis de nombreuses années, et actuellement entre NYU à New York et Facebook AI Research (FAIR) à Paris.
- Mais le paysage de la recherche en IA a considérablement changé : concurrence de la recherche privée, avec l'ouverture de centres de recherche fondamentale par les grands acteurs de l'IA, faible performance en termes de valorisation et de transfert vers l'industrie, que ce soit vers des startups ou vers des grands groupes et enfin fuite des cerveaux ...
- Les capacités françaises de formation universitaire et d'encadrement au niveau master ou doctorat sont devenues critiques.
 - Les filières master du domaine sont contraintes de refuser des étudiants brillants tant leurs salles de cours sont bondées, le cours de Stéphane Mallat au Collège de France se donne à guichet fermé, etc.
 - Les formations privées qui fleurissent sont de qualité inégale – et elles sont coupées de la recherche, ce qui, dans ce domaine en si rapide évolution qu'est l'IA aujourd'hui, n'est pas viable à long terme.
- La conséquence ? On assiste à une pénurie d'ingénieurs formés en IA alors même

- que le marché du travail a besoin d'ingénieurs de qualité.
 - Il n'y a que 300 000 chercheurs et praticiens de l'IA dans le monde, alors que la demande se chiffrerait par millions.
- L'IA impact l'ensemble des domaines de recherche et se heurte à la faiblesse de la recherche française en termes d'interfaces entre disciplines, puisqu'avoir de brillants chercheurs en mathématiques ou en informatique d'un côté et, par exemple, en physique-chimie ou en médecine de l'autre, ne suffit pas à faire prospérer les recherches interdisciplinaires.
- Le phénomène s'accroît quand on parle de science des données : alors que la numérisation des autres disciplines scientifiques posait jusqu'ici surtout des problèmes techniques, l'arrivée de l'IA définit des problématiques qui nécessitent une collaboration très étroite entre spécialistes.
- Pour tirer parti de la révolution de l'IA dans tous les domaines scientifiques, il est nécessaire de mettre en place une interdisciplinarité effective entre spécialistes en IA et chercheurs des autres disciplines.



- 1 **Fédérer les acteurs de la recherche autour** d'Instituts Interdisciplinaires d'Intelligence artificielle : **3IA**
- 2 **Des moyens de calcul** pour la recherche
- 3 Rendre plus **attractives les carrières** dans la **recherche publique**

- **Afin de mieux mailler le territoire et les domaines de recherche en intelligence artificielle, la mission avance trois propositions principales.**
 - **La création d'Instituts Interdisciplinaires d'IA (3IA)** au sein des établissements publics d'enseignement et de recherche. Ces instituts doivent être diversifiés géographiquement et thématiquement pour éviter les redondances en termes de domaines d'application ou de recherche pluridisciplinaire.
 - L'attribution **de moyens conséquents pour la recherche** en mettant en place d'une part un supercalculateur conçu spécifiquement pour les applications d'intelligence artificielle dédié à la recherche française et de leurs partenaires industriels dans le cadre de projets communs. D'autre part, il est nécessaire de déployer un cloud de l'intelligence artificielle qui aura vocation à être développé à un niveau européen pour impulser une dynamique industrielle européenne de l'IA.
 - La revalorisation **des carrières dans la recherche publique** en augmentant l'attractivité de la France pour les talents expatriés ou étrangers, en augmentant considérablement le nombre d'étudiants en master et en doctorat en IA, en revalorisant les salaires des chercheurs et en fluidifiant et en amplifiant les échanges académie-industrie.

3.

Anticiper et maîtriser les
impacts de l'IA sur le travail
et l'emploi





Avant de commencer cette partie dédiée à l'emploi, quelques chiffres clefs :

- 80 % des emplois seraient potentiellement automatisables à plus de 50 % selon le Conseil d'orientation pour l'emploi ;
- d'après le sondage médiamétrie que nous avons commandé, pour **94 % des Français, l'IA va transformer les emplois et 96 % d'entre eux considèrent qu'elle nécessite l'adaptation des réglementations sur le travail et la formation.**

Prendre la mesure de l'enjeu....



- **Le monde du travail est à l'aube de grandes transformations et n'y est encore que peu préparé.** Les incertitudes sur les conséquences du développement combiné de l'intelligence artificielle, de l'automatisation et de la robotique sont très importantes, notamment concernant les emplois détruits et créés.
- Néanmoins, **il apparaît de plus en plus certainement que la majorité des métiers et des organisations vont être très largement transformés** et qu'un nombre d'emplois vont être automatisés:
 - la première étude sur le sujet à avoir largement défrayé la chronique, celle de Frey et Osborne, de l'université d'Oxford, prévoyait une destruction de 47% des emplois d'ici une vingtaine d'années aux États-Unis.
 - En France, un cabinet de conseil, Roland Berger, a établi approximativement que 42% des emplois étaient menacés à un horizon similaire.
 - La dernière étude parue sur le sujet, produite par le Conseil d'orientation pour l'emploi, rattaché au Premier ministre, prévoyait, quant à elle, suivant une autre méthode, que 10% des emplois étaient menacés de disparition, mais que 50% des emplois seraient potentiellement automatisés à plus de 50%.
- Il est donc nécessaire de **prendre le problème à bras le corps et d'agir résolument, sans céder à la panique ni au fatalisme.**

Penser la complémentarité humain/IA



Pour voir la vidéo, cliquer sur l'image

Quelle forme va prendre la diffusion de l'IA dans mon organisation ?

Il est nécessaire de penser en amont les modes de complémentarité humain/IA.

Ce qui est le plus efficace à court terme en termes d'automatisation, ne l'est pas forcément à plus long terme, parce qu'il peut conduire à déqualifier les salariés (à les automatiser davantage). Le fait d'obéir exclusivement aux ordres d'une machine, de limiter sa marge d'erreur (et donc de manoeuvre) peut diminuer les qualités humaines que l'automatisation rend en fait nécessaire (capacités cognitives transversales // capacités créatives // compétences sociales et situationnelles // capacités de précision relatives à la perception et à la manipulation).

(Contre)-exemple des entrepôts Amazon ou des 2 manières d'automatiser les centres d'appels (1/ automatiser complètement 50% et laisser 50% des plus complexes aux salariés, qui évoluent donc ou 2/ guider 100% des salariés en leur disant quoi répondre sur un écran) ;

--> Au niveau de chaque entreprise, nécessité de penser les modes de complémentarité. Nécessité de développer un indice de bonne complémentarité au niveau des branches // national.

Organiser la transformation des organisations



Quelle méthode pour l'implémentation de l'IA ? Imposer ces transformations par le haut risque de conduire à de la violence vis-à-vis des salariés mais également à de l'inefficacité. Il est nécessaire de faire de ce sujet un élément plein et entier du dialogue social.

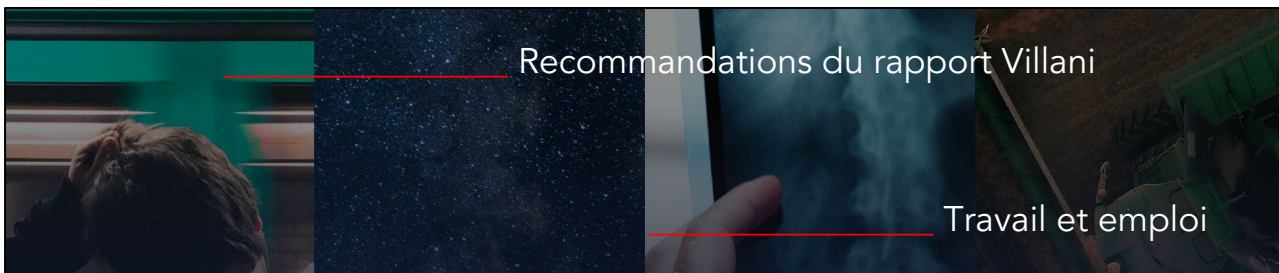
Les négociations sur la qualité de vie au travail et celles sur la GPEC (tous les 3 ans), peuvent être des canaux privilégiés de ce type de discussion.

De manière plus globale, **un chantier législatif relatif aux conditions de travail à l'heure de l'automatisation pourrait être lancé afin de prendre en compte les nouveaux risques :**

La législation concernant les conditions de travail est principalement adaptée aux modes de travail de l'ère industrielle et, malgré quelques avancées, notamment concernant le droit à la déconnexion, il semble qu'un ensemble de nouveaux risques et de nouvelles situations, liés au développement des technologies numériques dans l'entreprise, ne soient que difficilement pris en compte. Ainsi pour un grand nombre d'entreprises et d'individus l'encadrement des conditions de travail est trop rigide, voire inadapté aux nouvelles modalités du travail (mobilité, flexibilité...) tout en ne prenant pas en compte les risques spécifiques nouveaux. Le compte pénibilité est ainsi construit autour de la pénibilité industrielle. Il ne s'agit pas de remettre en cause cette approche, qui trouve à s'appliquer à un grand nombre de situations.

Néanmoins la nécessité d'encadrer des nouvelles situations de travail appelle l'ouverture d'un chantier législatif spécifique : obéissance exclusive aux instructions

d'une machine, impossibilité de discuter avec ses collègues sans passer par une interface machine... L'enjeu de la complémentarité devra être au cœur de cette redéfinition de l'encadrement des conditions de travail



- 1 Créer un **lab public** de la transformation du travail
- 2 Animer le **débat public** sur le **futur du travail**

Pour préparer au mieux cette transition, il est nécessaire de s'assurer que nos capacités d'anticipation d'observation de ces bouleversements soient pérennes, continues et surtout articulées avec des politiques publiques.

- Au sein de cette espace, les capacités prospectives, de prévisions macro-économiques et d'analyse des mutations des usages doivent être mises en lien avec des capacités d'expérimentation concrètes et articulées avec des actions à destination de certaines catégories de travailleurs.
- Placé au sein de l'appareil d'Etat, cette structure aurait pour rôle d'anticiper et d'expérimenter. Les fruits de ces travaux ont vocation à irriguer l'ensemble des politiques publiques de l'emploi et de la formation professionnelle.
- La démarche expérimentale pourra servir à amorcer des logiques différentes de celles qui sont actuellement en vigueur dans la formation professionnelle.
 - Les dispositifs actuels sont largement « à la main » des salariés, dans une logique de responsabilisation individuelle. Au vu du caractère potentiellement très rapide, voire exponentiel de ces transformations, il semble difficile, pour les dispositifs généraux existants, de répondre à l'ensemble des situations et de permettre à la fois la prise en compte des besoins de l'ensemble de la population et la nécessité d'agir de manière ciblée et urgente.
- C'est d'autant plus important que face à la transformation de leur emploi, les

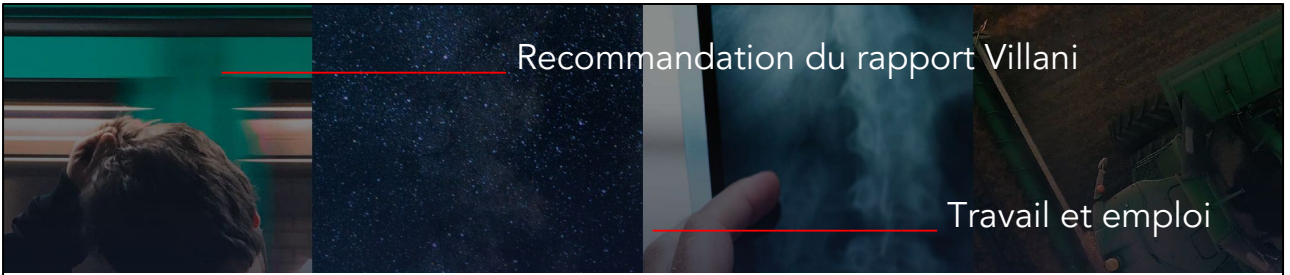
- parcours professionnels.
- À cet égard, des expérimentations pourraient être menées afin de construire des dispositifs qui ciblent certaines populations d'individus, dont les emplois sont considérés comme étant le plus à risque d'automatisation et pour lesquelles il sera complexe d'amorcer seules leur transition professionnelle. Il s'agit donc de rompre, en partie, avec la seule logique de responsabilisation de l'individu concernant sa propre transition professionnelle.
- **Le lab pourrait également avoir un rôle d'animation de réflexions ouvertes, au niveau national et international, sur la mutation du travail à l'heure de l'automatisation.**
- Ce rôle doit s'incarner, au-delà de l'aspect théorique, via l'instauration et/ou la mise en réseau de lieux de réflexion vivants sur l'avenir du travail, qui prennent la forme de laboratoires des nouveaux métiers et des usages de demain.

S'inspirer de l'Allemagne



En 2016, le gouvernement allemand a mis en ligne une **plateforme de débat public sur l'avenir du travail** : **www.arbeitsviernull.de**

- Ici, on peut s'inspirer de la plateforme Arbeitviernull en Allemagne
 - À l'occasion d'un congrès intitulé « Travail 4.0 » et organisé le 22 avril 2015 à Berlin, la ministre du Travail et des Affaires sociales (Andrea Nahles) a dévoilé un « Livre vert sur le Travail 4.0 » qui présente les principaux défis et questions soulevées par la révolution numérique en cours.
 - Ce document vise à servir de support à un autre débat sur l'avenir du monde du travail avec tous les acteurs concernés (décideurs économiques, politiques et sociaux, experts, citoyens).
 - Un site internet (www.arbeitsviernull.de) sert notamment de plateforme pour un dialogue large et ouvert avec toutes les parties prenantes



3

Repenser la **formation** à l'ère de l'IA

Le développement de l'intelligence artificielle nécessite une transformation des manières de former ainsi que des contenus de formation.

Loin de la logique adéquationniste qui consisterait à former des individus à occuper postes actuellement vacants, il convient de renforcer:

- l'apprentissage des compétences cognitives transversales pour apprendre à raisonner et à appréhender le monde de manière complexe
- la compétence matricielle dans un monde en perpétuelle évolution : la créativité.
 - L'apprentissage de la créativité suppose une modification profonde des modes d'enseignement : il ne s'agit pas de cantonner la créativité, entendue comme capacité de « développement personnel » à des matières « artistiques » ou à des enseignements annexes/d'ouverture, comme c'est souvent le cas aujourd'hui. C'est la manière d'enseigner les matières fondamentales qui doit être modifiée pour des pédagogies nouvelles qui soient orientées davantage vers le développement de l'esprit critique et de la coopération.

4

Développer la complémentarité Humains-machines

Si l'on part du principe que, pour la plupart des métiers, **les individus seront amenés à travailler en collaboration avec une machine, il est nécessaire de définir une complémentarité qui ne soit pas aliénante**, mais au contraire permette de développer les capacités proprement humaines (créativité, dextérité manuelle, capacité de résolution de problèmes...).

Cela pourra s'articuler de plusieurs manières.

1. **Tout d'abord au travers d'une transformation du dialogue social afin d'intégrer pleinement les enjeux numériques et développer un indice de bonne complémentarité.**
2. **De manière plus globale, un chantier législatif relatif aux conditions de travail à l'heure de l'automatisation pourrait être lancé afin de prendre en compte les nouveaux risques.**
3. **Une transformation de la formation initiale et continue devra être amorcée afin de favoriser les pédagogies expérimentales, à même de développer les compétences créatives qui deviennent de plus en plus cruciales.**

4 ●

L'IA au service d'une
économie
plus écologique



La consommation énergétique du numérique pourrait être multipliée par **10** en 2030

#AIforHumanity

AI

FOR

HUMANITY

- À l'heure où le réchauffement climatique est une certitude scientifique, la prise en compte des impacts environnementaux du développement des usages et services numériques est indispensable.
 - Or la révolution numérique et ses usagers occultent leurs impacts environnementaux.
- L'IA peut fournir des opportunités pour l'écologie : optimisation des ressources, préservation de notre environnement. Mais elle peut aussi entraîner des effets rebonds.
- Il faut donc aller au-delà d'un simple discours sur l'optimisation de l'utilisation de nos ressources et devenir le leader de la transition écologique intelligente.
 - Donner un sens à l'intelligence artificielle, c'est également penser sa soutenabilité, notamment écologique.
 - Il s'agit de penser une IA nativement écologique et de l'utiliser pour mieux penser l'impact de l'humain sur son environnement.
- **Il y a urgence :**
 - d'ici 2040, les besoins en espace de stockage au niveau mondial, fondamentalement corrélés au développement du numérique et de l'IA, risquent d'excéder la production disponible globale de silicium.
 - Rappelons également que la consommation énergétique du numérique pourrait atteindre 20 à 50 % de la consommation mondiale d'électricité en 2030 et pourrait ainsi être multipliée par 10 !

1

Favoriser la **convergence de la transition écologique et du développement de l'IA**

2

Penser une **IA moins consommatrice d'énergie**

- La France et l'Europe peuvent devenir le fer-de-lance de cette transition écologique intelligente, d'abord en **inscrivant le sujet à l'agenda international**.
 - Premier chantier : penser les impacts de l'IA sur la réalisation des objectifs de l'ONU sur le développement durable (ODD) : en quoi celle-ci en met certains sous contrainte, comment elle peut à l'inverse permettre d'en accélérer d'autres.
 - L'IA doit s'intégrer aux initiatives émergentes dans le cadre de l'Accord climat et du Pacte mondial pour l'environnement.
- Les acteurs des transitions numérique et écologique doivent se fédérer. Pour cela, il est nécessaire de **créer un lieu dédié à cette rencontre entre la recherche en IA et la recherche portant sur l'optimisation des ressources énergétiques**. Il s'agira de porter des projets à la croisée des sciences du vivant et de l'écologie, la recherche sur le climat et la météo.
- Le consommateur doit être acteur dans le verdissement de ces technologies. Notre mission propose ainsi la **mise en place d'une plateforme dédiée à la mesure de l'impact environnemental des solutions numériques intelligentes**. Cette plateforme devra s'accompagner d'un outil simple permettant à tout citoyen de prendre conscience de ces enjeux et de comparer l'impact environnemental de ces différents produits et services, logiciels et matériels.

- L'une des priorités est aussi de penser une IA moins consommatrice d'énergie. Nous devons penser l'innovation de rupture dans le domaine du semi-conducteur, l'une des briques matérielles de l'IA.
 - Par ailleurs, les pouvoirs publics doivent agir pour le verdissement de la chaîne de valeur et accompagner l'industrie du cloud européen dans le sens de sa transition écologique. Certains acteurs sont déjà exemplaires en termes d'optimisation de l'utilisation de l'énergie. Il est important de diffuser ces bonnes pratiques à l'ensemble du secteur. Un label pourrait être mis en place afin de valoriser les solutions les plus exemplaires.
 - Enfin, le verdissement de la chaîne de valeur de l'IA passera nécessairement par des architectures matérielles et logicielles ouvertes (*open hardware* et *open software*) qui, en plus d'être un facteur de confiance, peuvent permettre des économies d'énergie significatives et qui peuvent inspirer les initiatives en cours au niveau européen.

S'inspirer du recyclage de la chaleur produite



Quelques exemples de recyclage de la chaleur produite

Le datacenter de la banque Natixis à Marne-La-Vallée alimente en eau à 55° C le réseau de chauffage urbain d'une zone en cours d'urbanisation et le centre nautique intercommunal du Val d'Europe.

La société Stimergie a développé un mécanisme permettant de récupérer chaque année 1 MWh de chaleur par serveur, soit près de 60 % de la chaleur dégagée. Autrement dit la consommation des serveurs est réduite de plus de moitié. L'entreprise a signé plusieurs contrats d'installation ailleurs en France, notamment pour équiper un bâtiment résidentiel de 40 logements à Nantes, ainsi que la piscine de la Butte aux Cailles dans le 13e arrondissement de Paris.

3

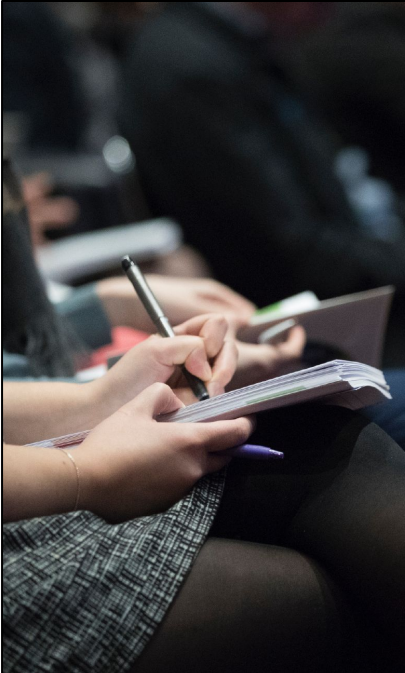
Libérer la donnée écologique

- Le développement d'une IA verte n'est possible qu'à condition de libérer la donnée écologique.
 - Pour développer des solutions d'IA au service de la transition écologique, il est ainsi primordial de mettre à la disposition de tous, chercheurs et entreprises européennes, et rapidement, à horizon 2019, les données publiques disponibles : météorologiques, agricoles, de transports, d'énergie, de biodiversité, de climat, de déchets, cadastrales, de diagnostic de performance énergétique... Pour les jeux de données les plus sensibles, l'ouverture pourrait se faire dans un périmètre précis, par exemple dans le cadre de défis sectoriels.
 - Il est également crucial de libérer la donnée privée, lorsque cela est nécessaire lorsqu'elle est utile.

A close-up photograph of a person's open palm holding a small, round, silver-colored compass. The hand is weathered and has some dirt on it. The compass has a black face with white markings and a red needle. The background is dark and out of focus.

5.

● Quelle éthique pour l'IA ?

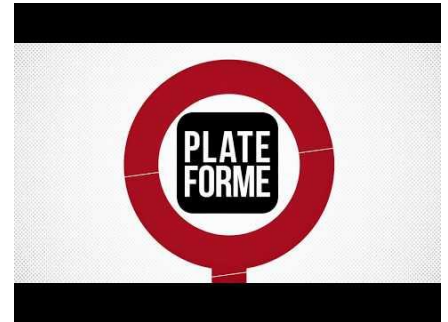


94% des sondés
estiment nécessaire que
**le développement de
l'IA fasse l'objet de
débats publics**

- L'intelligence artificielle affecte désormais tous les aspects de nos vies sociales. Sans toujours le savoir, nous interagissons quotidiennement avec des systèmes intelligents qui optimisent nos déplacements, créent nos playlists favorites ou protègent nos boîtes électroniques des courriels indésirables. Ils nous servent de manière invisible. C'est du moins l'objectif qu'on leur assigne : rendre nos vies meilleures, tâche par tâche.
- Les progrès récents de l'IA dans de nombreux domaines (voitures autonomes, reconnaissance d'images, assistants virtuels) et son influence croissante sur nos vies l'ont placée au centre du débat public. Ces dernières années de nombreuses voix se sont interrogées sur la capacité de l'IA à réellement œuvrer pour notre bien-être et sur les dispositions à prendre pour s'assurer que cela soit le cas.
- Ce débat a principalement pris la forme d'une large réflexion sur les enjeux éthiques liés au développement des technologies d'intelligence artificielle et plus largement des algorithmes. En différents endroits du monde, experts, régulateurs, universitaires, entrepreneurs et citoyens discutent et échangent régulièrement sur les effets indésirables, actuels ou potentiels de ceux-ci et les moyens de les atténuer.
 - D'ailleurs 94% des sondés estiment nécessaire que le développement de l'IA fasse l'objet de débats publics.
- Notre mission a cherché, humblement, à proposer quelques pistes permettant de poser les bases d'un cadre éthique pour le développement de l'IA et à faire vivre ce débat dans la société.

1

Ouvrir la boîte noire



- Une grande partie des considérations éthiques soulevées par l'IA tiennent à **l'opacité de ces technologies.**
- En dépit de leur performance accrue dans de nombreux domaines, de la traduction à la finance en passant par l'automobile, il est souvent très difficile d'expliquer leurs décisions de manière intelligible par le commun des mortels.
- De nombreux observateurs (chercheurs, activistes, pouvoirs publics) dénoncent régulièrement le fonctionnement en **boîte noire des systèmes algorithmiques**: **il est possible d'observer les données d'entrée (input), les données de sortie (output) mais dont on comprend mal leur fonctionnement interne.**
- **C'est particulièrement le cas pour les algorithmes d'IA car l'efficacité d'un algo de réseaux de neurones est inversement proportionnelle à son explicabilité.**

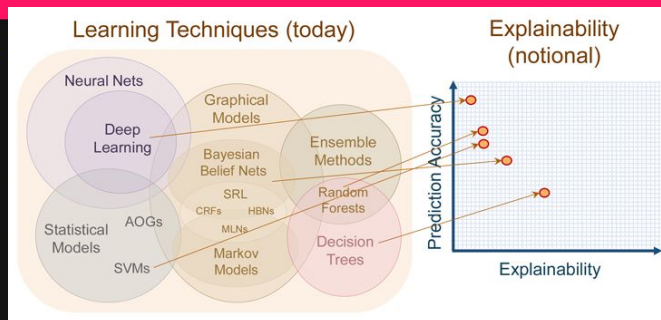
=> Pour [David Gunning de la Darpa](#), les progrès récents et rapides de ces technologies soulignent un paradoxe : l'exactitude de la prédiction de certaines techniques est inversement proportionnelle à leur explicabilité.

- L'utilisation d'arbres de décision par exemple permettent une meilleure explicabilité, mais ils se révèlent moins exacts dans leur capacité prédictive que l'apprentissage profond, qui donne de très bons résultats en terme de

- prédiction tout en restant obscur sur son fonctionnement. **Choisir entre le pourquoi et le quoi, risque d'être assez difficile.**
- c'est d'autant plus problématique que diverses études et controverses ont révélé les limites techniques et potentiels effets discriminants des algorithmes.
 - On pense notamment au cas des des algorithmes publicitaires de google qui proposent aux femmes des offres d'emploi moins rémunératrices que les hommes.

Le problème de l'explicabilité

BREAKING THE BLACK BOX



corps d'audit

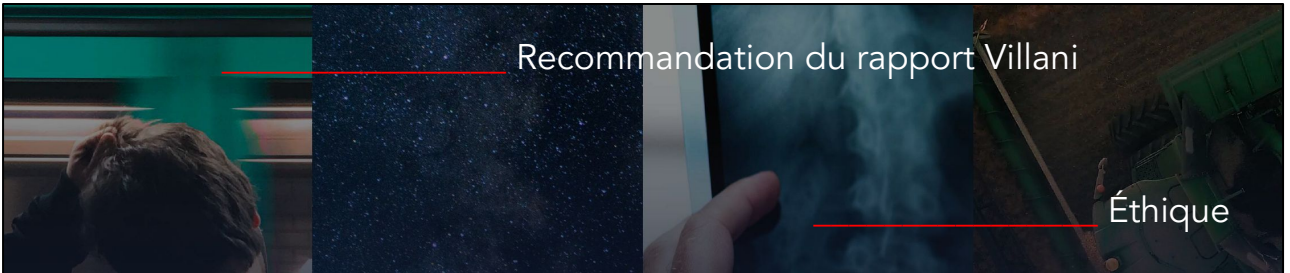
Recherche: explicabilité

Alors, comment parvenir à comprendre ce qu'il se passe dans les algorithmes d'IA ?

Plusieurs options existent :

- **Au niveau politique**, il est nécessaire de développer des **procédures, outils et méthodes permettant d'auditer ces systèmes** afin d'en évaluer la conformité à notre cadre juridique et éthique. C'est aussi nécessaire en cas de litige entre différentes parties mettant en cause les décisions prises par des systèmes d'IA.
 - Pour conduire cette mission d'audit, nous proposons:
 - **La création d'un corps d'audit des IA.**
 - Compétences : habilitation à procéder à des audits d'algorithmes et de bases de données sur pièce ainsi qu'à effectuer des tests par tout moyen requis.
 - Modes de saisines :
 1. Saisine par la justice dans le cadre d'enquêtes diligentées par des autorités administratives indépendantes;
 2. Saisine par le Défenseur des droits
 - **Le soutien au développement de l'évaluation citoyenne des IA.**

- Diverses options sont possibles:
 - Créer des fonctions de médiation publique dédiées à l'accompagnement de projets qui visent à mobiliser l'IA pour la compréhension des discriminations. Celles-ci travailleraient de manière étroite avec les autorités de régulation et les chercheurs.
 - Financer des résidences ou la constitution d'accélérateurs de projets (accompagnement scientifique, ingénierie, juridique, etc.) sur le modèle de Team up Turn, ProPublica ou EFF
 - Soutenir la rétro-ingénierie citoyenne par la mise à disposition de jeux de données et l'accompagnement d'experts en audit.
- **Au niveau de la recherche, il faut Investir massivement dans la recherche sur l'explicabilité de l'IA.** Ici, on peut s'inspirer du modèle de la DARPA (l'Agence américaine responsable des projets en recherche pour la Défense) qui a lancé un programme d'ampleur sur *Explainable AI* (explicabilité de l'IA) centré autour de trois axes:
 1. la production de modèles plus explicables
 2. la production d'interfaces utilisateurs plus intelligibles
 3. la compréhension des mécanismes cognitifs à l'œuvre pour produire une explication satisfaisante



2

Responsabiliser les développeurs et chercheurs en IA

Il ne faut pas uniquement développer des outils d'audit des algorithmes existants, il faut également à l'avenir considérer les enjeux éthiques dès la conception des algorithmes. A cette fin, il est nécessaire:

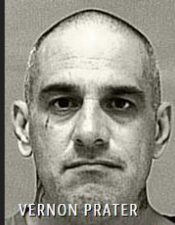
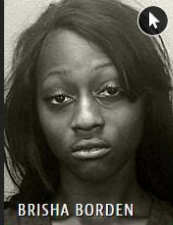
- **D'intégrer l'éthique dans la formation des ingénieurs et chercheurs en IA**
 - Cours d'éthique tout au long de la formation
 - Double cursus ingénierie - sciences sociales
- **D'Instaurer des étude d'impact sur les discriminations**
 - Les nouvelles règles européennes en vigueur imposent aux acteurs qui mettent en œuvre des traitements de données personnelles d'effectuer préalablement une étude d'impact quant à l'impact potentiel de ces activités sur les droits et les intérêts des concernés : **c'est l'étude d'impact sur la vie privée (privacy impact assessment, ou PIA).**
 - Selon la même logique et conformément aux lignes directrices du groupement européen des autorités de protection des données (G29), il serait bon d'introduire **un discrimination impact assessment (DIA)** pour amener les concepteurs d'IA à s'interroger sur les conséquences sociales des algorithmes qu'ils produisent.

Le cas particulier des algorithmes prédictifs



- De nombreuses institutions (**banques, police justice, armée**) commencent à **déployer des systèmes d'analyse prédictive à diverses fins**.
 - Dans le **secteur bancaire**, il s'agit par exemple de mieux prédire les probabilités de défaut de paiement des emprunteurs et ainsi d'optimiser l'allocation du crédit en fonction de leur niveau de risque respectif. C'est du moins toute la promesse du *big data*, renforcée par l'avènement de l'IA.
 - En australie, le **Service australien des douanes et de la protection des frontières (The Australian Customs and Border Protection Service)** a mis en place un système d'analyse du risque terroriste des passagers étrangers en partance pour l'Australie. Ce système, conçu par IBM, croise les données contenues dans les dossiers des passagers, ceux de services de renseignements australiens et les données sociales disponibles en ligne pour établir des profils de risque.
- En France, ces scénarios restent encore largement hypothétiques et le développement de ces initiatives n'en est qu'à un stade expérimental. Suivant l'exemple australien, **des services de police pourraient à l'avenir avoir recours à des algorithmes pour gérer le déploiement de leurs unités de patrouille et des armées utiliser des systèmes d'armes létales autonomes (SALA) sur des théâtres d'opérations extérieures**. De tels bouleversements, que ce soit dans la santé, la banque, l'assurance et plus particulièrement dans un cadre régalién, soulèvent d'importantes problématiques éthiques.

COMPAS (l'algorithme d'estimation des risques de récidive)

Two Petty Theft Arrests		Two Petty Theft Arrests	
			
VERNON PRATER	BRISHA BORDEN	VERNON PRATER	BRISHA BORDEN
Prior Offenses 2 armed robberies, 1 attempted armed robbery	Prior Offenses 4 juvenile misdemeanors	Prior Offenses 2 armed robberies, 1 attempted armed robbery	Prior Offenses 4 juvenile misdemeanors
Subsequent Offenses 1 grand theft	Subsequent Offenses None	Subsequent Offenses 1 grand theft	Subsequent Offenses None
LOW RISK 3	HIGH RISK 8	LOW RISK 3	HIGH RISK 8
Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.		Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.	

Cas d'études : L'enquête de Propublica

- En mai 2016, des journalistes de Propublica (journal d'investigation américain) ont révélé que l'algorithme d'estimation des risques de récidive COMPAS, développé par l'entreprise Northpointe et utilisé par la justice américaine, était raciste et inefficace.
 - En analysant les scores attribués aux détenus, il est apparu que cet algorithme surévaluait systématiquement le risque de récidive des détenus noirs-américains, à une fréquence deux fois plus élevée que celui des Américains blancs.
 - En plus, les détenus blancs étaient systématiquement représentés comme à faible risque, en contradiction avec leur taux réel de récidive.
- Cela signifie que cet algorithme a conduit à maintenir en détention des détenus noirs qui n'auraient probablement pas récidivé (faux positifs), tout en permettant la libération de probables récidivistes blancs (faux négatifs).
- En outre, l'enquête de Propublica nous rappelle que nous ne sommes pas tous égaux face à ces systèmes. Les données qu'ils exploitent, issues des bases de données de la police et de la justice, sont susceptibles d'être biaisées, reproduisant les biais normaux de la société. En les utilisant sans distance critique, on risque d'enfouir des discriminations dans le code et de les distribuer de manière systémique.

Que faire pour prévenir les dérives ?

- Informer les citoyens sur leurs droits
- Assurer une responsabilité à un être humain, selon des modalités préalablement définies
- Entamer un débat de société sur la place de l'automatisation dans les décisions humaines



1. En premier lieu, **il convient d'informer les citoyens sur leurs droits.**
 - En l'occurrence, le droit au recours effectif et, au-delà, celui du droit à l'explication sur les traitements de données, qui fondent une pratique de surveillance.
 - De ce point de vue, rappelons que la loi Informatique et Libertés posait dès 1978 le principe selon lequel «aucune décision administrative ou privée impliquant une appréciation sur un comportement humain ne peut avoir pour seul fondement un traitement automatisé d'informations donnant une définition du profil ou de la personnalité de l'intéressé», ajoutant que «toute personne a le droit de connaître et de contester les informations et les raisonnements utilisés dans les traitements automatisés dont les résultats lui sont opposés».
 - Ces dispositions, étendues par une loi du 6 août 2004, montrent que le législateur avait anticipé très tôt les dangers éventuels de ce type de dispositifs. Ces dispositions ont été reprises en substance dans l'article 22 du RGPD.
2. Il faut **s'assurer qu'à tout point du raisonnement, on puisse imputer une responsabilité à un être humain, selon des modalités préalablement définies.**
 - Plusieurs scénarios mériteraient d'être étudiés, qui vont d'une responsabilité individuelle (de la personne qui prend la décision, au concepteur de l'algorithme ou de la technologie en question) jusqu'à une responsabilité en cascade.

1. Entamer un débat de société sur la place de l'automatisation dans les décisions humaines

- Existent-ils des domaines où le jugement humain, aussi faillible soit-il, n'a pas vocation à être remplacé par la machine ? Si oui, il faut réfléchir dès à présent aux dispositions à prendre pour les protéger.

3

Organiser la gouvernance de l'éthique de l'IA



- La place de l'éthique dans le débat sur l'IA a pris aujourd'hui une importance telle **qu'il paraît nécessaire d'instaurer, dans un cadre institutionnel, un Comité consultatif national d'éthique pour les technologies numériques et l'intelligence artificielle**
- **Mission:**
 - **Cet organe serait chargé d'organiser le débat public, de façon lisible, construite et encadrée par la loi.**
 - Il devra parvenir à articuler des logiques de temps court, celui des enjeux économiques et industriels, en bonne interaction avec les comités sectoriels, tout en parvenant à s'en extraire pour penser le temps long.
 - **Les avis de ce comité, élaborés en toute indépendance, pourraient éclairer les choix technologiques des chercheurs, des acteurs économiques, industriels et de l'État.** Ses recommandations pourront servir de référence pour la résolution de dilemmes éthiques (par exemple sur le véhicule autonome) et donc servir de standard les développements en IA.



Pour une IA inclusive
et diverse



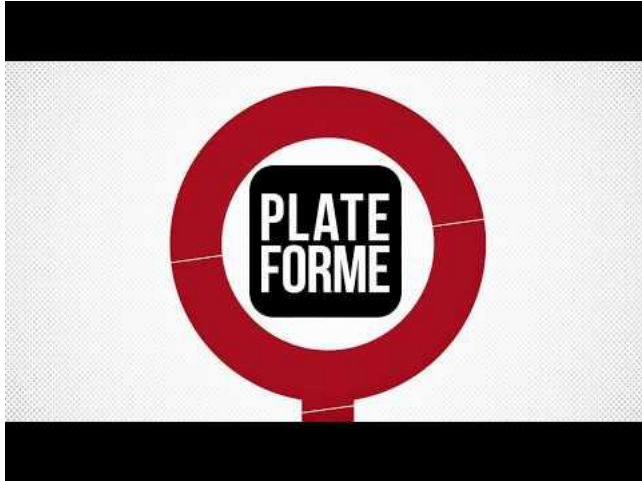
Mixité et diversité

La sous-représentation des femmes en IA est problématique alors même que le manque de diversité peut conduire les algorithmes à reproduire des biais.

- L'IA ne peut pas être une nouvelle machine à exclure.
- Au contraire, elle ouvre de formidables opportunités de création de valeur et de développement de nos sociétés et des individus. Ces opportunités doivent bénéficier à tous.
- À tous, et en premier lieu à toutes.
 - Rappelons tout d'abord que les filières scientifiques et techniques se féminisent (certes lentement). Le numérique, quant à lui, fait figure d'exception : la parité entre les hommes et les femmes est loin d'y être acquise. Les femmes représentent à peine 33% des personnes du secteur du numérique (12% seulement, si l'on écarte les fonctions transversales et supports).
 - Ce manque de diversité peut conduire les algorithmes à reproduire des biais cognitifs — souvent inconscients — dans la conception des programmes, l'analyse des données et l'interprétation des résultats.
 - Les illustrations sont nombreuses : on constate par exemple que de nombreux programmes associent des mots tels que « programmation » avec « homme » et « tâches ménagères » avec « femme ». En 2015, lorsque vous tapiez « CEO » (PDG) dans un moteur de recherche aux États-Unis, il fallait aller jusqu'à la 96e photo pour y trouver une femme... sans compter qu'il s'agissait d'une image de *Barbie CEO* en tailleur minijupe ! Aujourd'hui

- encore, lorsqu'on tape ce mot dans un moteur de recherche, il en ressort très majoritairement des photos d'hommes, alors que près d'un tiers des PDG américains sont des femmes.
- L'un des grands défis de l'IA consiste donc à parvenir à une meilleure représentativité de nos sociétés. Pour relever ce défi, il nous faut nous appuyer sur l'ensemble de nos talents. Cette mobilisation collective doit être la plus ouverte possible. Chacun doit avoir les mêmes chances de participer à la recherche, au développement et la création de valeur en IA. Tout l'enjeu consiste donc à décroïsonner et à distribuer les capacités d'innover.

Un exemple de cas : CEO Barbie



- Les illustrations de biais genrés sont nombreuses : on constate par exemple que de nombreux programmes associent des mots tels que « programmation » avec « homme » et « tâches ménagères » avec « femme ». En 2015, lorsque vous tapiez « CEO » (PDG) dans un moteur de recherche aux États-Unis, il fallait aller jusqu'à la 96e photo pour y trouver une femme... sans compter qu'il s'agissait d'une image de *Barbie CEO* en tailleur minijupe ! Aujourd'hui encore, lorsqu'on tape ce mot dans un moteur de recherche, il en ressort très majoritairement des photos d'hommes, alors que près d'un tiers des PDG américains sont des femmes.
- L'un des grands défis de l'IA consiste donc à parvenir à une meilleure représentativité de nos sociétés. Pour relever ce défi, il nous faut nous appuyer sur l'ensemble de nos talents. Cette mobilisation collective doit être la plus ouverte possible. Chacun doit avoir les mêmes chances de participer à la recherche, au développement et la création de valeur en IA. Tout l'enjeu consiste donc à décloisonner et à distribuer les capacités d'innover.

1

Porter une politique de féminisation audacieuse du secteur du numérique et de l'IA

- Si l'éducation à l'égalité et au numérique est une condition préalable et essentielle, la mixité pourrait être atteinte avec une politique incitative visant à atteindre un seuil de 40% d'étudiantes dans les filières du numérique (classes préparatoires et filières des grandes écoles et des universités) d'ici 2025.
 - Ainsi, si les établissements parviennent rapidement à une proportion de 40% de filles, ces derniers pourraient être récompensés par une labellisation ou une subvention.
 - Un exemple : la Grande École du Numérique a pour but de favoriser la mixité des publics dans le secteur du numérique et de faire bénéficier les femmes des opportunités offertes par le numérique. Les formations labellisées ont ainsi pour objectif d'accueillir au moins 30% de femmes au sein de leurs promotions d'apprenants.
- L'ensemble des actions en faveur de la diversité dans les entreprises du numérique pourraient par ailleurs être portées par une action nationale en faveur de la mixité et de la diversité dans la technologie avec l'alimentation, notamment, d'une base de données nationale permettant d'objectiver les inégalités entre les femmes et les hommes au travail et de fonds dédiés à soutenir la diversité dans l'IA.

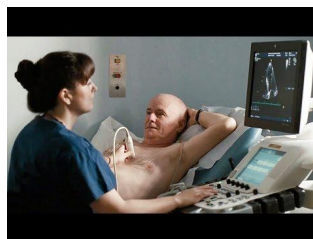


Inclusion : il faut s'assurer que personne ne soit mis de côté.

- Plus généralement, face à l'évolutivité des technologies et des usages liés à l'IA, notre société est tenue à un devoir de **réflexivité** et de **vigilance collective**. Notamment vis-à-vis des populations fragiles et des publics déjà exclus du numérique, pour qui l'IA peut représenter un danger encore plus grand. Le développement de l'IA peut tout autant emporter la promesse d'une société meilleure, plus juste et efficace, que le risque d'une hyperconcentration de la valeur au profit d'une petite élite numérique.
- En préalable à une politique d'inclusion, l'éducation du plus grand nombre aux principes de l'IA et des algorithmes est primordiale. L'introduction d'une matière dédiée aux humanités numériques, récemment annoncée par le ministre de l'Éducation nationale, pourrait servir de support à ces enseignements.
- En matière d'IA, la politique d'inclusion doit donc revêtir un double objectif :
 - d'une part s'assurer que le développement de ces technologies ne contribue pas à accroître les inégalités sociales et économiques ;
 - d'autre parts'appuyer sur l'IA pour effectivement les réduire.
- Plutôt que de fragiliser nos trajectoires individuelles et nos systèmes de solidarités, l'IA doit prioritairement nous aider à **activer nos droits fondamentaux, augmenter le lien social et renforcer les solidarités**.

2

Développer la médiation numérique et l'innovation sociale pour que l'IA bénéficie à tous



- Face à l'ampleur des transformations à venir par l'IA, il est de notre responsabilité collective de s'assurer que personne ne soit mis de côté.
- Pour que chacun puisse véritablement bénéficier des avancées de l'IA, nos procédures d'accès aux droits doivent évoluer, et nos capacités de médiation considérablement se renforcer.
- Notre mission propose donc de mettre en place un système automatisé d'aide à la gestion des démarches administratives qui vise à améliorer la connaissance pour le grand public des règles administratives et de leur application à une situation personnelle.
 - En complément, de nouvelles capacités de médiation doivent être déployées pour accompagner les personnes qui en ont besoin, en lien avec les réseaux du prendre soin présents sur le territoire.
- Enfin, il est important que la puissance publique soutienne le développement d'initiatives basées sur l'IA dans les champs sociaux. À l'heure actuelle, les capacités d'innovation par l'IA restent très concentrées au sein d'un petit nombre d'entreprises. À l'exception de la santé, les champs sociaux reçoivent une part minoritaire des investissements privés.
- Cette structuration de l'écosystème d'innovation en IA a des conséquences sur la vitesse des progrès réalisés dans les champs sociaux. Afin de redistribuer ces capacités d'innovation, la puissance publique pourrait lancer des programmes spécifiques pour accompagner l'innovation d'IA en matière sociale, et outiller les

- acteurs sociaux afin qu'ils puissent bénéficier des avancées liées à l'IA.

La stratégie numérique de la France —

Suite à la remise du rapport Villani, le 29 mars au Collège de France, le Président de la République a décliné une **stratégie sur l'intelligence artificielle** reposant sur **4 piliers** (détail en annexe) :



- Renforcer l'écosystème français de l'IA en mettant l'**accent sur la recherche et l'expérimentation**
- Engager une **politique résolue d'ouverture des données**
- Développer la présence française en IA sur **quelques secteurs clés**
- Poser un **cadre éthique**

La vision — La révolution de l'IA pour l'Etat

L'intelligence artificielle est entrée, depuis quelques années, dans une nouvelle ère. Cette transformation radicale est à la fois une opportunité majeure et une responsabilité immense pour l'État qui doit se réinventer.

À l'heure où la course à l'IA s'intensifie, où plusieurs pays, dont les Émirats Arabes Unis et le Royaume-Unis se dotent d'entités publiques spécialisées en IA, l'annonce par le PR de la stratégie gouvernementale en la matière est l'occasion de créer, au sein de la DINSIC, un véritable espace de pilotage, de coordination, d'expérimentation et de réflexion dans la mise en œuvre de cette stratégie. **Objectifs :**

- **Un État exemplaire et moderne :** dans le cadre du plan Action publique 2022, la stratégie de l'Etat en matière d'IA devra permettre à la fois de renforcer l'efficacité et la qualité des services publics. Il s'agira de fluidifier les relations entre les citoyens et des entreprises avec l'Etat, dans un cadre qui permet l'inclusion de tous et le respect strict de règles éthiques.
- **Un État moteur de la transformation de la société et de l'économie :** la politique publique pour l'intelligence artificielle devra permettre de faire émerger des projets à large échelle, en s'appuyant sur les ressources dont dispose l'Etat, notamment via le partage de données.
- **Un État plateforme :** les écosystèmes de l'IA devront être étroitement associés à la construction des services publics de demain. Startups d'Etat et entrepreneurs d'intérêt général permettront d'intégrer des solutions d'innovation et des principes d'amélioration au travail de l'administration, selon des finalités communes.



Le rapport est disponible sur :

AI

FOR

HUMANITY

.fr

L'INTELLIGENCE ARTIFICIELLE
AU SERVICE DE L'HUMAIN.

- L'intégralité du rapport ainsi que la stratégie gouvernementale sont consultables sur le site aiforhumanity.fr. Vous pouvez également y revivre la journée de conférence du 29 mars dernier.